

**EXTENDING CHEMICAL COMPLENATION TO BACTERIA AND
FURTHERING NUCLEAR RECEPTOR BASED PROTEIN ENGINEERING
AND DRUG DISCOVERY**

A Thesis
Presented to
The Academic Faculty

by

Kenyetta A Johnson

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in the
School of Chemistry and Biochemistry

Georgia Institute of Technology

August 2009

**EXTENDING CHEMICAL COMPLEMENTATION TO BACTERIA AND
FURTHERING NUCLEAR RECEPTOR BASED PROTEIN ENGINEERING
AND DRUG DISCOVERY**

Approved by:

Dr. Donald Doyle, Advisor
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Bridgette Barry
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Andreas Bommarius
School of Chemical & Biomolecular
Engineering
Georgia Institute of Technology

Dr. Joe Ledoux
School of Biomedical Engineering
Georgia Institute of Technology

Dr. Ichiro Matsumura
School of Chemistry
Emory University

Dr. Adegboyega Oyelere
School of Chemistry and Biochemistry
Georgia Institute of Technology

Date Approved: May 14, 2009

The ultimate measure of a man is not where he stands in moments of comfort and convenience, but where he stands at times of challenge and controversy.

- Martin Luther King Jr., Strength to Love 1963

I dedicate this work to my mother, Vera Consuella Shury. I am all that I am because of you! I thank you for loving me, for encouraging me, and for helping me to be the best that I can be.

ACKNOWLEDGEMENTS

Graduate school has been a process in which amazing people were put in my path to teach me more than I ever knew I would learn, which helped to not only develop me as a scientist but as an individual as well. I would first like to thank my advisor, Dr. Donald Doyle. There have been many challenges in both of our lives both professionally and personally, but he has always provided an environment for me to learn and grow. I thank him for his guidance and support throughout the years and for teaching me to think on my own and to develop my own ideas. I would also like to give a special thanks to Dr. Bahareh Azizi who has had such a large impact in my development into the scientist that I have become. I cannot imagine how this experience would have been without her. I would like to thank her for her friendship, for pushing me, for teaching me all I know about molecular biology, but most importantly for setting an example for me! I will always cherish the support, guidance, and friendship that you have given me both professionally and personally.

I would also like to thank all of my committee members Dr. Bridgette Barry, Dr. Andreas Bommarius, Dr. Joseph Ledoux, Dr. Ichiro Matsumura, and Dr. Adegboyega Oyelere for all of their support throughout the years, their guidance and encouragement, letters of recommendation, and the tireless editing of my thesis. I would like to give special thanks to Dr. Oyelere and his lab for their support and guidance with the dual inhibiting compounds project. I thank all of my lab members, past and present, who have become more than friends to me but my family. I would like to thank you Dr. Lauren Schwimmer and Dr. Terry Watt for their in depth scientific conversations, advice, and

ssistance in many experiments. I would like to give a special thanks to Dr. Priyanka Rohatgi for encouraging me to join the Doyle lab and for teaching me cell culture. She always helped to make my time in the lab so much more “interesting!” I am grateful for her amazing professional advice and support and for being such a wonderful friend. I would like to thank Eileen Chang for her meticulous help with so many experiments and helping me to take my mind off of research when needed. I would like to thank my current lab members, Hilda Castillo, Anna Duraj-Thatte, James Kratzer, Amanda Ousley, Michael Rood, Hally Shaffer, and Jennifer Taylor, for helping me so much especially during the writing of my thesis by assisting me in experiments and making figures, but most importantly for their support. I would like to give an extra thanks to Hilda Castillo and Jennifer Taylor for burning the midnight oil with me and for supporting me during my “breakdowns” and spurts of “hyperness”.

I would like to thank the Bommarius lab members, Janna Billy-Blum, Dr. James Broering, Dr. Javier Chaparro-Riggers, Dr. Bernard Loo, Dr. Karen Polizzi, and Dr. Tracey Thaler for their camaraderie, support, and the many memories. I would like to give special thanks to Dr. Eduardo Vasquez-Figueroa not only for your support with research but also for his inspiring drive, positive attitude, and continuing counsel and friendship. On a personal basis I would also like to give a special thanks to Banafsheh Azizi for being a wonderful friend and always being there for me, listening to me, encouraging me, and always giving me a positive outlook.

Finally, I would like to thank all of my family for their support. I especially want to thank my mother, Vera C. Shury, for being my strength and for teaching me the

importance of education and how to work hard and strive for the best. I would not have accomplished what I have without her love, support, and guidance.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	v
LIST OF TABLES	x
LIST OF FIGURES	xi
ABBREVIATIONS	xv
SUMMARY	xvii
CHAPTER 1 NUCLEAR RECEPTORS	1
1.1 Nuclear Receptor Structure	1
1.2 Nuclear Receptor Function	5
1.3 Retinoid X Receptor	10
1.4 Literature Cited	14
CHAPTER 2 THE POWER OF GENETIC SELECTION	23
2.1 Genetic Selection	23
2.2 Yeast Two-Hybrid Systems	25
2.3 Chemical Complementation	30
2.4 Bacterial Two-Hybrid Systems	35
2.5 Literature Cited	41
CHAPTER 3 BACTERIAL CHEMICAL COMPLEMENTATION: ELIMINATING BACKGROUND	47
3.3 Bacterial Chemical Complementation: Bacteria vs. Yeast	47
3.2 Developing Bacterial Chemical Complementation	49
3.3 Results of Eliminating Bacterial Chemical Complementation Background	57
3.3.1 Effects of Expression Vectors on Background Activation	58
3.3.2 Effect of Promoters on Background	58
3.3.3 Effects of Integrating the Reporter on Background Growth	61
3.4 Materials and Methods	63
3.5 Literature Cited	67
CHAPTER 4 BACTERIAL CHEMICAL COMPLEMENTATION: LIGAND DEPENDENT ACTIVATION	70
4.1 Optimizing Bacterial Chemical Complementation for Ligand Activation	70
4.1.1 Enhancing the Interaction Between GBD and Gal1p	70

4.1.2 Optimizing the Various Components of the BCC system	76
4.2 Enhancing Heterologous Protein Expression and Stability	82
4.2.1 Analysis of Fusion Proteins Solubility	85
4.2.2 Enhancing Heterologous Expression and Folding with Chaperones	87
4.2.3 Heterologous Expression and Folding with Osmolytes	93
4.3 Summary of Bacterial Chemical Complementation	97
4.4 Future of Bacterial Chemical Complementation	99
4.5 Materials and Methods	100
4.6 Literature Cited	105
 CHAPTER 5 LIGAND-RECEPTOR CHARGE REVERSAL VIA GENETIC ENGINEERING	 113
5.1 Engineering Nuclear Receptors and Orthogonal Ligand Receptor Pairs	113
5.2 Motivation for Engineering Ligand-Receptor Charge Reversal	117
5.3 Retinoic Acid Receptor	120
5.4 Results of Ligand-Receptor Charge Reversal	123
5.4.1 Design of Engineered Receptor Library for Charge Reversal	123
5.4.2 Effects of Amide Based Retinoids on Charge Reversal Mutants	127
5.4.3 Effects of Amine Based Retinoids on Charge Reversal Mutants	130
5.5 Summary and Future Work of Engineering Ligand-Receptor Charge Reversal	142
5.6 Materials and Methods	144
5.7 Literature Cited	145
 CHAPTER 6 DUAL INHIBITING COMPOUNDS FOR BREAST CANCER THERAPEUTICS	 151
6.1 Estrogen Receptor and Breast Cancer	151
6.2 Motive for Dual Inhibiting Compounds as Breast Cancer Therapeutics	156
6.3 Using Chemical Complementation to Assess Activity of Dual Inhibiting Compounds towards Estrogen Receptor	161
6.4 Whole Cell Proliferation Assays in Estrogen Receptor Positive and Negative cells	166
6.4.1 Effects of Dual Inhibiting Compounds in ER α Positive Breast Cancer Cells	168
6.4.2 Effects of Dual Inhibiting Compounds in ER α Negative Breast Cancer Cells	174
6.4.3 Conclusions of Effects of Dual Inhibiting Compounds in Breast Cancer Cells	176
6.5 Summary of Dual Inhibiting Compounds as Breast Cancer Therapeutics	177
6.6 Future Work	179
6.8 Materials and Methods	181
6.7 Literature Cited	183

LIST OF TABLES

	Page
Table 5.1: Library of Site Directed Charge Reversal Variants	126
Table 5.2: Charge Reversal Variants Activation by Neutral Amide Ligands	129
Table 5.3: Charge Reversal Variants Activation by Positively Charged Amine Ligands	134
Table 6.1: Half Maximal Inhibitory Concentrations (IC ₅₀) of Compounds in ER ⁺ and ER ⁻ Cancer Cell Lines	169

LIST OF FIGURES

	Page
Figure 1.1: Nuclear Receptor Structure	4
Figure 1.2: Topology and Structural Framework of Nuclear Receptor Ligand Binding Domain	4
Figure 1.3: Nuclear Receptor Function: Activation and Repression	8
Figure 1.4: Effects of Ligands and Coregulators on Helix 12 Position	8
Figure 1.5: Retinoid X Receptor Binding Pocket Residues and their interaction with 9cRA	13
Figure 2.1: General Yeast Two-Hybrid System	27
Figure 2.2: Genetic Complementation	32
Figure 2.3: Chemical Complementation	34
Figure 2.4: Negative Chemical Complementation	34
Figure 2.5: Repression Based Bacterial Two-Hybrid System (B2H)	37
Figure 2.6: Activation Based Bacterial Two-Hybrid System	39
Figure 2.7: Wood's Bacterial One-Hybrid Construct	39
Figure 3.1: Bacterial two-hybrid System (BTHS)	50
Figure 3.2: Bacterial Chemical Complementation (BCC)	50
Figure 3.3: Applications of Bacterial Chemical Complementation	53
Figure 3.4: Bacterial Chemical Complementation Expression and Reporter Vectors	56
Figure 3.5: Initial Results of Bacterial Chemical Complementation	56
Figure 3.6: Effects of Expression and Reporter Vectors on Background Growth Nuclear Receptor Structure	59
Figure 3.7: Effect of BCC Promoters on Background Growth	59
Figure 3.8: Conjugation and Integration of the Reporter Construct	62

into the KJIC Strain

Figure 3.9:	Effects of Reporter Construct Integration on Background Growth	64
Figure 4.1:	Construction of the Ligand-Independent Control	71
Figure 4.2:	Initial Results of Ligand Independent Control	72
Figure 4.3:	Effects of Gal4 Response Elements on Ligand Independent Control	74
Figure 4.4:	Bacterial Chemical Complementation with Four Gal4 Response Elements	75
Figure 4.5:	Bacterial Chemical Complementation with Various Nuclear Receptor-Ligand Pairs and Coactivators	75
Figure 4.6:	Nuclear Receptor-Ligand Pairs	79
Figure 4.7:	Results of Various Nuclear Receptor-Ligand Pairs in Bacterial Chemical Complementation	81
Figure 4.8:	RNA Polymerase Core Enzyme and Subunits	81
Figure 4.9:	Bacterial Chemical Complementation with the Sigma70 RNA Polymerase Subunit	83
Figure 4.10:	Evaluating Coactivator Fragments in Yeast Chemical Complementation	83
Figure 4.11:	Results of GBDRXRLBD and Coactivator Fragments in Bacterial Chemical Complementation	86
Figure 4.12:	Expression of Bacterial Chemical Complementation Fusion Proteins	88
Figure 4.13:	Model for Chaperone Assisted Protein Folding in <i>E. coli</i>	90
Figure 4.14:	Effects of Chaperones on Bacterial Chemical Complementation	92
Figure 4.15:	Bacterial Chemical Complementation in Proline Enhanced Media	95
Figure 4.16:	Bacterial Chemical Complementation in High Concentrations of Proline Enhanced Media	98
Figure 5.1:	Orthogonal Ligand-Receptor Pairs	114
Figure 5.2:	Crystal Structures of RAR γ and RAR α	122
Figure 5.3:	Retinoid Ligands	126

Figure 5.4:	Retinoic Acid Receptor α Chemical Complementation	128
Figure 5.5:	(A) Activation of R274E;K236E;K229E Variant with atRA and EthAmD RA (B) Activation of RARwt with Amide Ligands	131
Figure 5.6:	Ligand-Receptor Activation Profiles of K236E-EthAmD RA and RARwt- atRA	133
Figure 5.7:	RARwt and R274E;K236E;K229E Activation with atRA	133
Figure 5.8:	(A) Stabilization of atRA by Positively Charged Residues in RAR wt (B) Activation of RARwt and R278E with atRA	136
Figure 5.9:	Activation of Single Variants by atRA	138
Figure 5.10:	Activation of RARwt, R274E;K236E;S289, and K236E;S289D;K229E Variants by EthAmN RA	138
Figure 5.11:	Positively Charged Potential Surrounding the Carboxylate of atRA	140
Figure 5.12:	“Up” and “Down” Conformations of K236 in RAR γ	140
Figure 5.13:	Binding of atRA into Binding Pocket of RARwt in “Up” and “Down” Conformations of K236	141
Figure 5.14:	Binding of EthAmN RA into Binding Pocket of RARwt in “Up” and “Down” Conformations of K236	143
Figure 6.1:	(A) Crystal Structure of Estrogen Receptor Alpha bound to Estradiol and (B) Estrogen Receptor Agonists	152
Figure 6.2:	HDAC inhibitors, Estrogens, and Anti-estrogen effects on Estrogen Signaling	159
Figure 6.3:	Estrogen Receptor Based Chemical Complementation	162
Figure 6.4:	Dual Inhibiting Compounds and Tamoxifen and SAHA	162
Figure 6.5:	Growth Dose Responses of Estrogen Receptor alpha with Estradiol, DY-001-137, DY-001-138, and DY-001-148	164
Figure 6.6:	Estrogen Receptor Antagonist Chemical Complementation	165
Figure 6.7:	Inhibition of Estrogen Receptor Alpha by Tamoxifen, DY-001-137, DY-001-138, and DY-001-148	165

Figure 6.8:	(A) Agonist and (B) Antagonist β -galactosidase Assays using Chemical Complementation	167
Figure 6.9:	Estrogen Receptor (A) alpha and (B) beta Antagonist β -galactosidase Dose Response Curves with Tamoxifen, DY-001-137, DY-001-138, and DY-001-148	167
Figure 6.10:	Effects of Tamoxifen, DY-001-137, DY-001-138, and DY-001-148 in MCF-7 Cells Using MTS Cell Proliferation Assay	170
Figure 6.11:	Proliferation of MCF-7 Cells in the presence of SAHA and DY-001-148	172
Figure 6.12:	Proliferation of MCF-7 Cells in the presence of DY-001-148 and Equal amounts of Tam and SAHA	173
Figure 6.13:	Proliferation of MDA-MB-231 Cells in the presence of Tamoxifen, DY-001-137, DY-001-138, and DY-001-148	175
Figure 6.14:	Proliferation of MDA-MB-231 Cells in the presence of SAHA, DY-001-137, DY-001-138, and DY-001-148	178
Figure 6.15:	Proliferation of MDA-MB-231 Cells in the presence of DY-001-148 and Equal amounts of Tam and SAHA	180

ABBREVIATIONS

3AT	3-amino-1,2,4-triazole
9cRA	9- <i>cis</i> retinoic acid
ACTR	Human activator for thyroid and retinoid receptors
ADE	Adenine
atRA	All trans retinoic acid
B2H	Bacterial two hybrid system
BCC	Bacterial chemical complementation
CC	Chemical complementation
DBD	DNA binding domain
EC ₅₀	Half maximal effective concentration
ER	Estrogen receptor
GAD	Gal4 activation domain
GBD	Gal4 DNA binding domain
HDAC	Histone deacetylases
HDACi	Histone deacetylases inhibitor
HIS	Histidine
HRE	Hormone response element
IC ₅₀	Half maximal inhibitory concentration
LBD	Ligand binding domain
LXR	Human liver X receptor
NR	Nuclear receptor
OLRP	Orthogonal ligand-receptor pair

ONPG	<i>o</i> -nitrophenyl β -D-galactopyranoside
PGC1a	Mouse peroxisome proliferator-activated receptor γ coactivator-1
PXR	Human pregnane X receptor
RAR	Human retinoic acid Receptor
RNAP	RNA polymerase
RXR	Human retinoid X receptor
SERM	Selective estrogen receptor modulator
SRC-1	Steroid receptor coactivator 1
Wt	Wild-type
Y2H	Yeast two hybrid system

SUMMARY

Nuclear receptors (NRs) are modular ligand-activated transcription factors that control a broad range of physiological processes by regulating the expression of essential genes involved in cell physiology, differentiation and metabolism. These receptors have been implicated in a number of diseases, such as cancer and diabetes. Due to their profound role in development and disease progression, much emphasis is being put forth into nuclear receptor based drug discovery. Furthermore, due to the modularity of these proteins, there is also an emphasis towards engineering these receptors to bind novel small molecules, creating orthogonal ligand receptor pairs, in which the receptor only binds the synthetic ligand and the synthetic ligand is unable to bind the natural receptor. These orthogonal ligand receptor pairs have potential for applications in areas such as gene therapy and for the creation of biosensors.

Chemical Complementation (CC) is a genetic selection based assay in yeast that was developed to aid in the discovery of these orthogonal ligand receptor pairs. This system exploits the modularity and functionality of nuclear receptors. This three-hybrid yeast assay has the nuclear receptor ligand binding domain fused to the Gal4 DNA binding domain (GBD), and a nuclear receptor coactivator fused to the Gal4 activation domain (GAD). Upon binding of ligand, the nuclear receptor recruits the coactivator:GAD fusion protein, resulting in expression of a selection marker. CC is a powerful tool for discovering and developing orthogonal ligand receptor pairs. Due to several advantages, to include faster growth times and higher transformation efficiencies, we have attempted to extend chemical complementation from yeast to *E. coli*.

The bacterial chemical complementation system (BCC) was designed based on a bacterial two hybrid system in which the alpha subunit of RNA polymerase is fused to a nuclear receptor coactivator and the GBD is fused to a nuclear receptor and expressed in *E. coli*. The Gal4 DNA binding domain binds its response element upstream of an essential gene. Ideally, upon binding of the appropriate ligand to the NR ligand binding domain (LBD), the LBD undergoes a conformational change, recruits the coactivator-alpha RNA polymerase fusion protein, and activates transcription of an essential gene. A new bacterial strain was engineered in which a Gal4 response element controls the expression of the *HIS3* gene in the bacterial strain. Once this strain was created and the background growth for this strain was reduced using various concentrations of 3-aminotriazole (a competitive inhibitor of the *HIS3* gene), we were able to produce activation from our ligand independent control system. However, bacterial chemical complementation did not produce ligand dependent activation.

To obtain ligand dependent activation we investigated various aspects of the BCC system, including the use of various nuclear receptor-ligand pairs, various coactivator proteins, and smaller fragments of both the nuclear receptor and coactivator proteins. The inability to obtain ligand dependent activation using BCC is most likely due to inability to attain heterologous protein expression of eukaryotic proteins in *E. coli*. Protein chaperones and osmolytes were used to try and enhance heterologous protein expression, but no ligand dependent activation was achieved.

In a second project designed to further NR based protein engineering and drug discovery, chemical complementation in *S. cerevisiae* was used to evaluate a library of charge reversal variants. These variants were rationally designed in an attempt to gain a

better understanding of nuclear receptor function and structure and to produce an orthogonal ligand receptor pair. A library of retinoic acid receptor (RAR α) variants were developed to alter the binding selectivity of the receptor from the natural negatively charged ligand, all-trans retinoic acid (atRA), to positively charged retinoid ligands. Single, double, and triple variants were constructed based on five residues in the binding pocket of RAR γ known to stabilize the carboxylate of atRA. Multiple variants were evaluated via chemical complementation with diverse activation profiles with the various amine based retinoids and atRA. We were able to engineer two triple variants capable of activating with the ethyl amine retinoid but not the natural atRA ligand. However these variants do not activate with the ethyl amine retinoid as well as RAR α does. With the data obtained from evaluating the tolerability of mutations by RAR α , further developments towards engineering an enhanced ligand-receptor pair capable of activating with higher affinities of the amine based retinoids should be possible.

In a third project CC was utilized to characterize tamoxifen and histone deacetylase inhibitor based dual inhibiting compounds as breast cancer therapeutics. The compounds were assessed for their ability to inhibit estrogen receptor (ER) activation and thus decrease cell proliferation associated with breast cancer. The dual inhibiting compounds are composed of a covalently linked tamoxifen based moiety and a histone deacetylase inhibitor (HDACi) based moiety. Both tamoxifen alone and HDACi alone have been found to decrease proliferation in breast cancer cell lines via ER and histone deacetylases, respectively. However, covalently linking the two moieties can potentially create dual inhibiting compounds that act on various stages of the cell cycle to produce a more potent and effective drug to treat breast cancer. Several dual inhibiting compounds

were found to decrease the activation of ER, by its natural ligand estradiol, better than tamoxifen alone. Additionally, when tested in both ER positive and negative breast cancer cells, the compounds were found to decrease proliferation better than tamoxifen alone or the HDACi alone. We have also established that covalent linkage of SAHA and tamoxifen enhances their anti-proliferative effects in MCF-7 cells. Overall, these compounds can have a profound impact as a potential therapeutics as an alternative method for breast cancer treatment.

CHAPTER 1

NUCLEAR RECEPTORS

1.1 Nuclear Receptor Structure

Nuclear receptors (NRs) are ligand-activated transcription factors that control a broad range of physiological processes, including cell differentiation, proliferation, and maintaining homeostasis [1-4]. NRs control these processes by regulating the expression of genes that encode key enzymes, transporters, and other proteins involved in practically every facet of mammalian physiology. With the sequencing of the human genome, 48 human nuclear receptors have been identified [6]. These proteins mediate the actions of steroid and non-steroid hormones, as well as fatty acids, vitamins, and other small molecules in the body [4, 7-10]. The modulation of transcription by NRs occurs through various mechanisms, including activation and repression [4, 11, 12].

Nuclear receptors have been found to be involved in a vast array of reproductive, proliferative, and metabolic diseases including cancer, diabetes, and infertility [13-16]. The involvement of these transcription factors in disease makes them excellent drug targets. Nuclear receptor ligands account for approximately 10% of all commonly prescribed drugs [2, 17, 18]. Some of these drugs include: the estrogen receptor ligand tamoxifen, which targets breast cancer [16], the glucocorticoid receptor ligand dexamethasone, which targets inflammatory disease [14], and the peroxisome proliferator-activated receptor ligand thiazolidiones, which targets type II diabetes [13]. As a result of their implication in various diseases and their ligand controllability, there has been a large push towards research not only to discover ligands for numerous NRs but also to better understand their structure and function [19-29].

Nuclear Receptor Structure

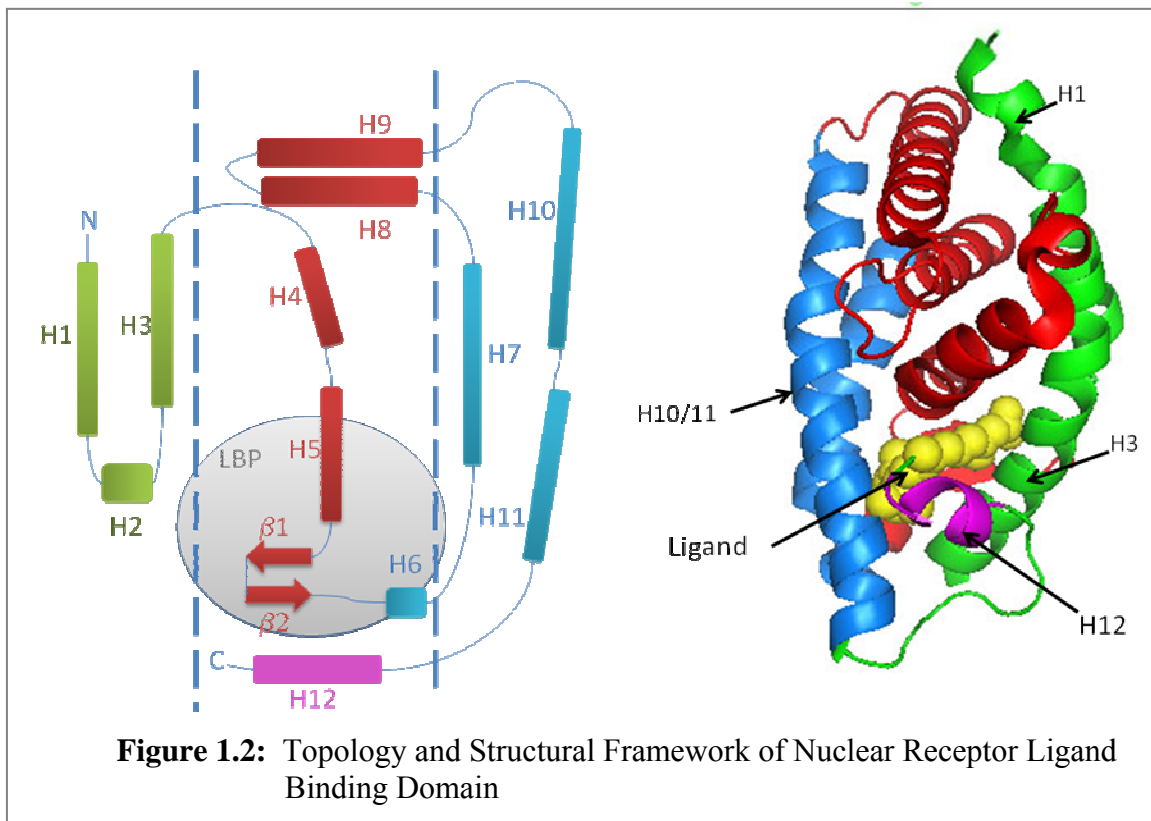
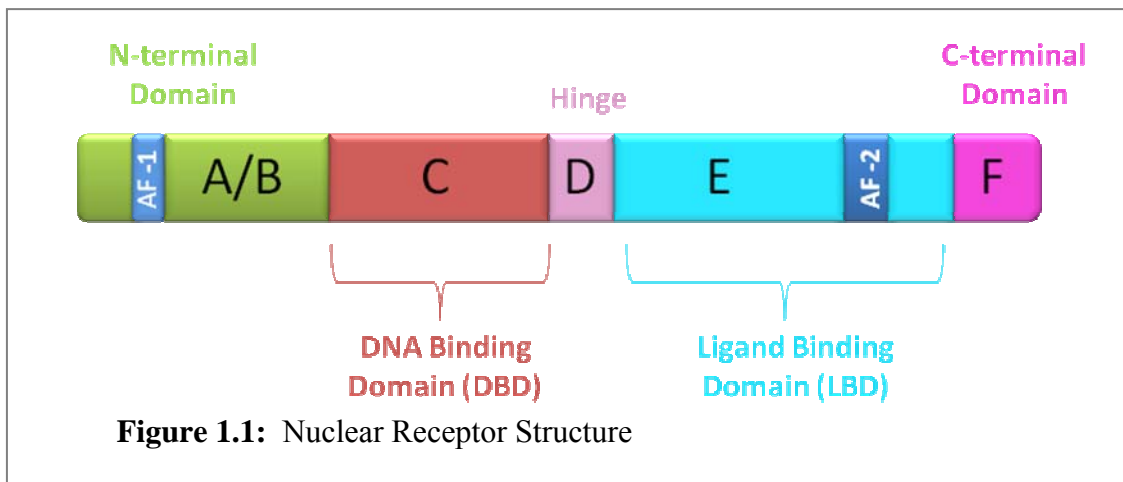
Nuclear receptors are modular proteins, containing five to six conserved domains designated A to F, which include an N-terminal domain, DNA-binding domain (DBD), the hinge region, ligand binding domain (LBD), and a C-terminal domain respectively (Figure 1.1). The N-terminal A/B domains and the hinge regions are less conserved than the other domains, whereas the C-terminal F domain, which is contiguous with the E domain, is not present in all nuclear receptors. All three of these domains have weakly understood functions. Although poorly understood, the N-terminal A/B domain has been found to contain an autonomous ligand independent transcriptional activation function, referred to as AF-1. Additionally, A/B domains vary in length and sequence, have been shown to be post-translational modification targets, and have the ability to interact with co-regulators [30-36]. The D domain serves as a hinge between the DBD and LBD allowing for their rotation, conferring structural flexibility and minimizing steric hindrance between the DBD and LBD [37]. Little is known about the F domain, but is implicated in the regulation of ligand binding and dimerization [37].

The DBD and LBD (regions C and E respectively) are the most studied and most highly conserved domains and are able to function independently. The DBD is responsible for binding DNA sequences called hormone response elements (HRE) [38]. Typical HREs contain two hexa-nucleotide motifs (generally AGGTCA) separated by several nucleotides. DBD-DNA binding specificities are achieved by this spacing as well as the orientation of the half sites (direct-, indirect-, everted-, or inverted-repeats). The DBD is known to have a helix-loop-helix structure. Crystallographic studies as well

as nuclear magnetic resonance studies reveal that this domain is made of two cysteine-rich zinc finger motifs, two α -helices, and a COOH extension [39, 40]. The DBD core contains several sequence elements, A, D, P, and T boxes, shown to not only contribute to binding specificities, but also to define the DBD's dimerization interface and contacts with both the DNA backbone and flanking HRE residues [41]. The P and D boxes contain the two conserved zinc-fingers and determine sequence specificity and half-site spacing, respectively. The C-terminal region of this domain contains a nuclear localization sequence responsible for nuclear entry and has also been suggested to contain a nuclear export signal [42].

The LBD is functionally intricate and highly structured, with roles in dimerization, ligand recognition, as well as cofactor interactions [43]. Nuclear receptor LBDs have a common fold consisting of 12 α -helices (H1-H12) and a β -turn arranged as an antiparallel α -helical sandwich in a three layer structure [44] (Figure 1.2). In most NRs, the first 3 helices make up one face of the LBD. H6, H7, and H10 correspond to the other face, whereas H4, H5, the β -turn, H8, and H9 compose the central layer [37]. H12 corresponds to the activation function helix, AF-2 domain, known to mediate ligand-dependent activation as well as coactivator recruitment. In addition to coactivator interaction the LBD can interact with corepressors as well.

In spite of the LBDs conserved fold, the ligand-binding pocket varies greatly in size and ligand specificity. Various hydrophobic, polar, and charged residues line the ligand binding pocket of different receptors. Polar residues are mainly found at the deep end of the pocket acting as ligand anchoring points or playing roles in ligand positioning and specificity. Overall, ligand specificity is determined by amino acid residues as well



as the shape and size of the binding pocket, leading to the various activation profiles and functions observed between NRs and their subtypes [45].

1.2 Nuclear Receptor Function

Nuclear receptors regulate transcription using several distinct mechanisms, including DNA recognition, transcriptional activation, and repression. DNA binding is an essential step in NR function. These receptors interact via specific DNA sequences, HREs, positioned upstream of target gene promoters. The sequence of the HRE half sites, the spacing between the half sites, and the orientation of the half sites directs sequence specific recognition of DNA by NRs. These proteins bind DNA as either monomers (such as the steroidogenic factor-1, SF1) [46], homodimers (such as the estrogen receptor, ER) [40], or heterodimers with the retinoid X receptor (RXR) (such as the retinoic acid receptor, RAR) [47]. These dimerization patterns dictate the response element specificity based from both the spacing between half-sites as well as response element orientation. For example, various RXR heterodimers bind to direct repeats (DRs) of the two half-sites with between one and five base pairs between them, termed DR1 through DR5 [47-50]. The type of HRE dictates which target genes are regulated by what type of NR.

Transcriptional regulation is controlled by the LBD and its ability to interact with coregulators. These coregulators modulate transcription by participating in chromatin remodeling. In the absence of ligand, NRs are either present in the cytoplasm, such as steroid receptors [51], or in the nucleus constitutively bound to their respective HREs [52] and complexed with corepressors, which recruit histone deacetylases (HDACs) (Figure 1.3 A). The corepressors that directly interact with NRs include the NR

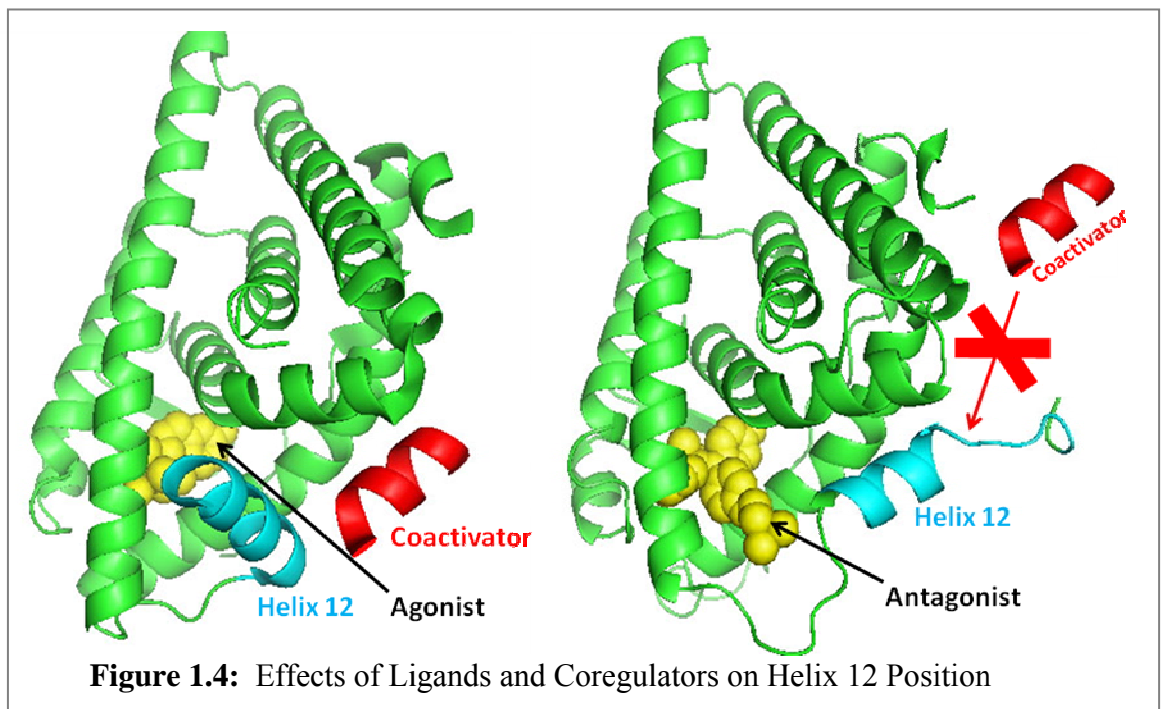
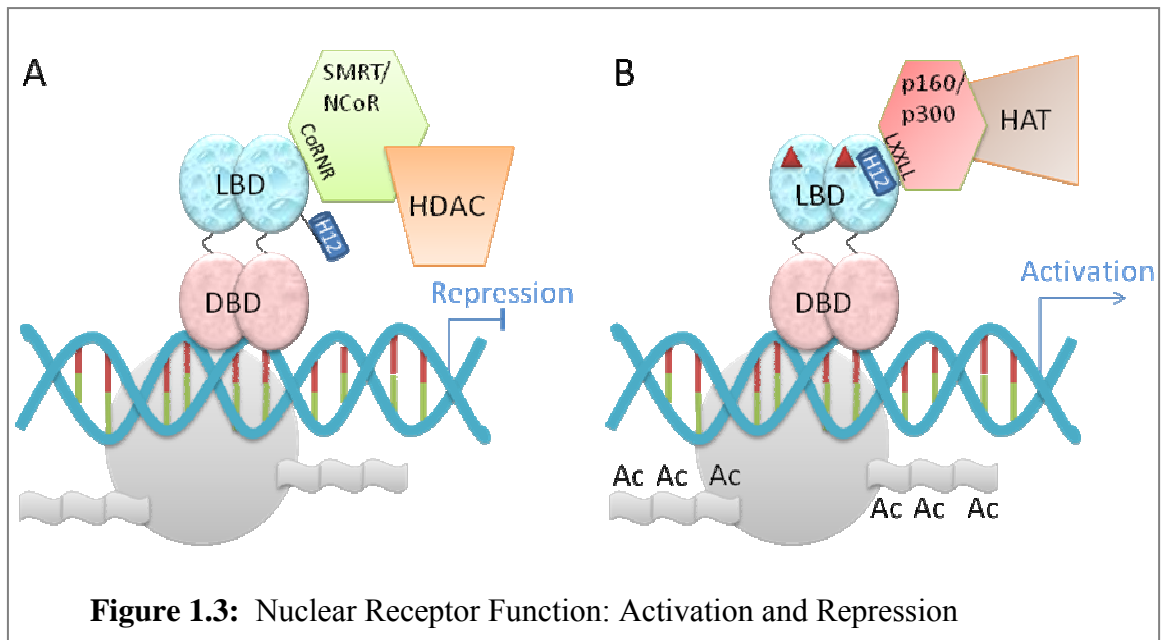
corepressor (NCoR) and the silencing mediator of retinoid and thyroid receptors (SMRT). These corepressors interact with NRs via a corepressor NR box (CoRNR box) [53]. This CoRNR box contains a LXXXIXXXI/L motif responsible for interacting with a hydrophobic groove on the surfaces of helices 3 through 5 on the NR LBD [54]. Either in the absence of ligand or when an antagonist ligand is bound, these NR bound corepressors, NCoR and SMRT, are involved in recruiting chromatin remodeling proteins, known as HDACs. HDACs are enzymes that catalyze the removal of acetyl groups from lysines in histones causing a tighter association between the histones and DNA and preventing the access required for RNA polymerases to initiate transcription and express the gene of interest [55-58].

The first step in transcriptional activation is the binding of an agonist ligand to the LBD. This action activates the NR by inducing a conformational change in the LBD, mediated by helix 12 (AF-2 domain). For cytoplasmic NRs, ligand binding additionally induces translocation from the cytoplasm to the nucleus, initiating regulation of transcription. The mechanism of this conformational change in the AF-2 domain has been observed by comparing the crystal structures of ligand bound (holo) and unbound (apo) receptors [12, 44, 59, 60]. Generally, upon agonist binding, H12 repositions against the core of the LBD, sealing the ligand binding pocket, contributing to the hydrophobicity of the pocket and stabilizing ligand binding while creating a surface for coactivator interactions. In addition to the repositioning of H12, ligand binding also induces conformational changes that affect other helices in the LBD [12, 61]. In particular, the repositioning of H12 disrupts the hydrophobic groove between H3-H5, known to interact with corepressor proteins, not only producing a more compact structure, but also

producing surfaces on the LBD which favor coactivator binding and disrupt corepressor binding [62] (Figure 1.4).

Ligand induced coactivator recruitment allows diverse families of coactivators to bind to NR LBDs. The p160 and p300 family of coactivators are the first to be recruited to ligand activated NRs [63-65]. The p160 family of coactivators include the steroid receptor coactivator 1 (SRC-1) and the activator of thyroid and retinoic acid receptor (ACTR). These coactivators have been found to bind approximately the same hydrophobic groove as the CoRNR box, making contacts on the surface of H3, H4, and H12 through a highly conserved NR box containing LXXLL motifs [66, 67] (as opposed to the larger LXXXIXXXI/L motifs of the CoRNR box). The position of H12 is essential for coactivator binding (Figure 1.4). With the repositioning of H12, the three turn helix motif of the CoRNR box no longer fits into the hydrophobic groove of the LBD. This repositioning is now ideal for the two turn helix motif of coactivator LXXLL motifs to bind. Once coactivator is bound, H12 forms a charge clamp with H3, locking the coactivator in place [54, 68].

In addition to interacting with the NRs, the p160 coactivators described above also interact with cAMP response element binding protein (CBP), also a member of the p160 family, and p300 family coactivators. Both p160 and p300 coactivators act as histone acetyltransferases (HATs) [65]. HATs are enzymes which catalyze the addition of acetyl groups to lysines in histones, relaxing the tight association between the histones and DNA, allowing chromatin expansion and the access required for RNA



polymerases to initiate transcription and express the gene of interest (Figure 1.3 B).

In addition to transcriptional activation, nuclear receptors are also able to exhibit transcriptional repression. Like activation, repression can be mediated by ligand binding of an antagonist small molecule as opposed to an agonist. NR agonists are small molecules that bind and upregulate gene expression as described previously. NR antagonists have the opposite effect, they bind causing a down regulation of gene expression [69, 70]. Antagonists bind the NR and block transcription activation through competitively binding to the same binding site as the agonist, but with a structure that sterically displaces H12, preventing coactivator recruitment and enabling corepressor binding (Figure 1.4). Antagonists do not allow the agonist small molecule to bind resulting in no conformational change of the LBD and thus no coactivator recruitment or transcriptional activation. Consequently, antagonist based transcriptional repression is mediated by a conformation of the receptor that prevents coactivator association and promotes association with the same corepressors involved in unliganded NR repression, NCoR and SMRT, repressing transcription via HDACs [71].

Ligands that bind NRs and cause activation or repression are not always true agonist or antagonists. Some ligands can also be inverse agonists or partial agonists. Inverse agonists are synthetic compounds that reduce the basal levels of NR activation that occur in the absence of an agonist. Whereas agonist and antagonist have distinct stable positioning of the AF-2 domains, inverse agonist cause the AF-2 domains to adopt an alternative position that does not obstruct the hydrophobic groove formed by H3 and H4 [72]. Partial agonists/antagonists are compounds that are potent but exhibit poor

efficacy in comparison to true agonist or antagonists. This decreased efficacy is a result of decreased stabilization of the AF-2 domain [73, 74].

With the continuing research to better understand NR function and structure, several developments into ligand discovery and design have been made. Understanding NR structure has provided insights into protein engineering as a tool to design NR ligand selectivity. This knowledge can not only enhance drug potency, by allowing researchers to designing specific ligands that activate receptors more effectively and at lower concentrations than wild-type ligands, but this knowledge can also aid in manipulating drugs to produce fewer side effects, by creating orthogonal ligands only capable of binding the receptor of interest. Drug discovery targeted towards NRs has aided in the discovery and development of both agonists and antagonists as well as selective nuclear receptor modulators (SNuRMs). SNuRMs are drugs that display an agonist response in some tissues and an antagonistic response in others. A common example of a SNuRM is the breast cancer therapeutic agent tamoxifen, which is generally an antagonist in breast cells, therefore used as a breast cancer treatment [75], but acts as an agonist in bone tissue and the endometrium, aiding in the prevention of osteoporosis but increasing uterine cancer risks [76, 77].

1.3 Retinoid X Receptor

The retinoid X receptor, RXR, plays important roles in numerous fundamental biological processes including reproduction, cellular differentiation, bone development, hematopoiesis, and pattern formation during embryogenesis [18]. Structurally, RXR is similar to other nuclear receptors containing the variable A/B domain, DBD, hinge region, and LBD. However, this nuclear receptor is also unique in comparison to other

nuclear receptors. RXR is the only nuclear receptor known to form heterodimers with other member of the nuclear receptor superfamily [78-80]. Although, much research has been done to investigate RXR, there are still many unknown or controversial issues regarding RXR, including its true biological role and the existence of an actual endogenous ligand. Despite its controversy, RXR is an essential NR due to its necessity as a heterodimerization partner with other NRs involved in various critical biological pathways. As a result of its implication in these various processes, RXR has been identified as a potential target for cancer therapy as well as metabolic diseases [81, 82].

There are three RXR subtype: RXR α , β , and γ . RXR α is expressed predominately in the liver, kidney, epidermis and intestine, and is the major RXR found in the skin [83-85]. RXR β is expressed widely and can be detected in all tissues [83, 85-87], whereas RXR γ is constrained to the muscle, pituitary, and parts of the brain [83, 85, 88, 89]. All three subtypes are heterodimerization partners with members of the subfamily 1 NRs, which include the retinoic acid receptor and the thyroid hormone receptor [17].

Interaction of these subfamily 1 receptors with RXR is not only required for their function, but RXR also increases their DNA binding efficiency [90]. These RXR heterodimers bind an arrangement of response elements containing direct repeats with between 1 and 5 base pairs separating them (DR1-DR5) [78]. Due to RXRs wide distribution and the diversity of HRE it is able to bind, the amount of RXRs potential target genes is vast, further supporting the importance of RXR in various biological processes. RXR has also been implicated to be involved in RXR specific signaling as well, due to its *in vitro* ability to form homodimers that bind DR1 response elements and

activate genes [90, 91]. Nevertheless, the physiological role of RXR homodimers has yet to be discovered.

RXR heterodimerization partners are classified as functionally “permissive” or “nonpermissive” [92]. Permissive heterodimers are synergistically activated by both RXR agonists as well as agonists of the dimer partner either together or individually. Nonpermissive heterodimers cannot be activated by RXR agonists but only by agonists of the dimer partner [93]. RXR in nonpermissive heterodimers is still able to bind its agonist, however nonpermissive partners inhibit its activation. Interestingly, permissive partners have been found to bind ligands with a low affinity whereas nonpermissive partners tend to bind their ligands with a high affinity.

RXR was originally described as an orphan receptor, a receptor whose endogenous ligand has not yet been identified [84]. Only after finding that it was activated by very high concentrations of all-*trans*-retinoic acid (at-RA), an isomer of at-RA, 9-*cis*-retinoic acid (9cRA), was discovered to activate all three subtypes of RXR, RXR α , RXR β , and RXR γ [83]. However, much controversy surrounds the existence of 9cRA as a truly endogenous ligand. Although 9cRA has been shown to be present in developing embryos and biosynthetic enzymes of 9cRA have been identified, 9cRA has not been clearly detected in mammalian cells and as a result cannot be concluded as the natural ligand for RXR [94, 95]. Phytanic acid and docosahexaenoic acid have been proposed as natural ligands for RXR [96-98], but neither have been proven to be true natural RXR ligands, leaving this issue controversial and in need of further investigations.

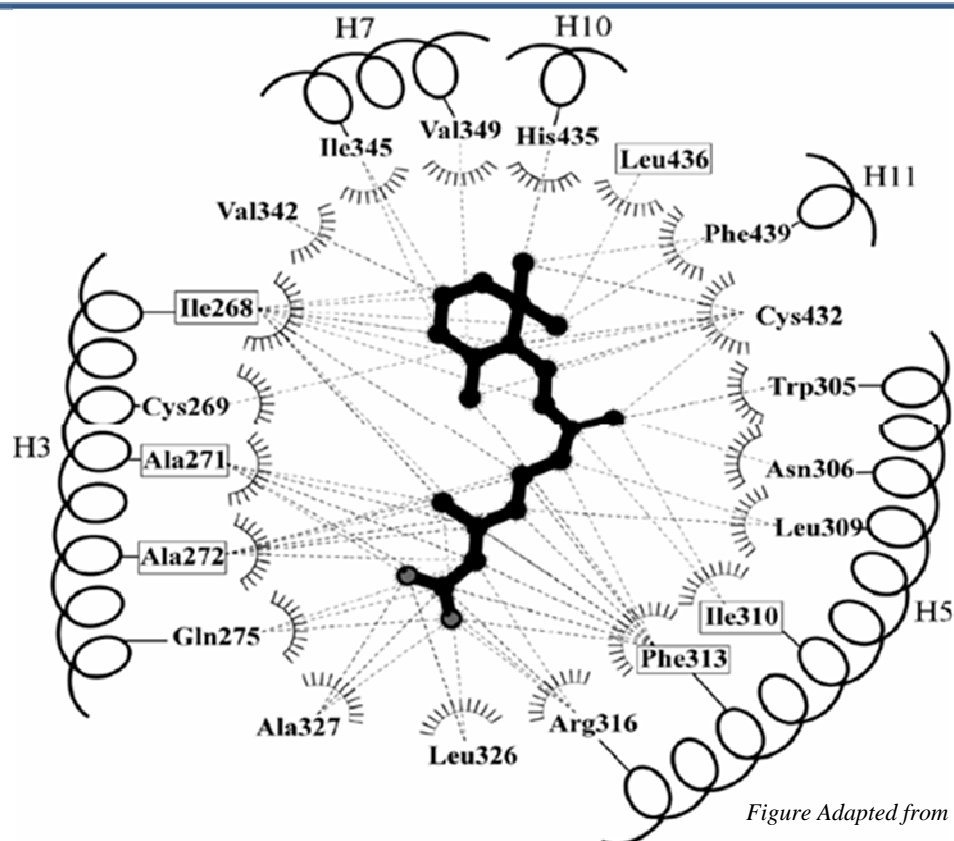


Figure Adapted from [5]

Figure 1.5: Retinoid X Receptor Binding Pocket Residues and their interaction with 9cRA

Although 9cRA has been defined as a RXR agonist, it is not a RXR selective compound because it also activates the retinoic acid receptors (RAR) with high affinity. Resultantly, synthetic compounds (rexinoids) that only activate RXR have become important in order to determine the role of these receptors. Crystal structures of both RXR and RAR bound to various ligands have provided insight into the LBPs of both receptors [99, 100], thus providing a means to design specific ligands able to discriminate between the two receptors. The crystal structures have also provided insight into what residues within the LBP are essential for ligand binding. RXR bound to 9cRA was one of the first nuclear receptor crystal structures to be solved, allowing researchers to determine the key interaction between the receptor and the ligand [101] (Figure 1.5). Structural analysis of NRs has allowed researchers to manipulate the LBP to bind unnatural ligands [22, 23].

The ability of NRs to bind small molecules and regulate the expression of critical genes involved in various reproductive, proliferative, and metabolic diseases makes these proteins excellent targets for drug discovery [13-16]. As a result, significant research has gone into engineering NRs to respond to non-natural ligands. To engineer these receptors, proteins libraries need to be created and assayed for binding and function. A powerful assay for analyzing NRs is through the power of genetic selection.

1.5 Literature Cited

1. Chambon P (2005) The nuclear receptor superfamily: a personal retrospect on the first two decades. *Mol Endocrinol* **19**, 1418-1428.
2. Gronemeyer H, Gustafsson JA & Laudet V (2004) Principles for modulation of the nuclear receptor superfamily. *Nat Rev Drug Discov* **3**, 950-964.

3. Katzenellenbogen JA & Katzenellenbogen BS (1996) Nuclear hormone receptors: ligand-activated regulators of transcription and diverse cell responses. *Chem Biol* **3**, 529-536.
4. Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schutz G, Umesono K, Blumberg B, Kastner P, Mark M, Chambon P & Evans RM (1995) The nuclear receptor superfamily: the second decade. *Cell* **83**, 835-839.
5. Rohatgi P (2006) Engineering protein molecular switches to regulate gene expression with small molecules. In, pp. 1 v. (various pagings). Georgia Institute of Technology, Atlanta, Ga.
6. Robinson-Rechavi M, Carpentier AS, Duffraisse M & Laudet V (2001) How many nuclear hormone receptors are there in the human genome? *Trends Genet* **17**, 554-556.
7. Beato M, Herrlich P & Schutz G (1995) Steroid hormone receptors: many actors in search of a plot. *Cell* **83**, 851-857.
8. Evans RM (1988) The steroid and thyroid hormone receptor superfamily. *Science* **240**, 889-895.
9. Gronemeyer H (1991) Transcription activation by estrogen and progesterone receptors. *Annu Rev Genet* **25**, 89-123.
10. Tsai MJ & O'Malley BW (1994) Molecular mechanisms of action of steroid/thyroid receptor superfamily members. *Annu Rev Biochem* **63**, 451-486.
11. Novac N & Heinzl T (2004) Nuclear receptors: overview and classification. *Curr Drug Targets Inflamm Allergy* **3**, 335-346.
12. Bourguet W, Germain P & Gronemeyer H (2000) Nuclear receptor ligand-binding domains: three-dimensional structures, molecular interactions and pharmacological implications. *Trends Pharmacol Sci* **21**, 381-388.
13. Kersten S, Desvergne B & Wahli W (2000) Roles of PPARs in health and disease. *Nature* **405**, 421-424.
14. Laue L, Kawai S, Brandon DD, Brightwell D, Barnes K, Knazek RA, Loriaux DL & Chrousos GP (1988) Receptor-mediated effects of glucocorticoids on inflammation: enhancement of the inflammatory response with a glucocorticoid antagonist. *J Steroid Biochem* **29**, 591-598.
15. Lin RJ, Egan DA & Evans RM (1999) Molecular genetics of acute promyelocytic leukemia. *Trends Genet* **15**, 179-184.
16. Shao W & Brown M (2004) Advances in estrogen receptor biology: prospects for improvements in targeted breast cancer therapy. *Breast Cancer Res* **6**, 39-52.

17. Laudet V GH (2002) *The Nuclear Receptor FactBook* Academic Press, London.
18. McDonnell DP, Vegeto E & Gleeson MA (1993) Nuclear hormone receptors as targets for new drug discovery. *Biotechnology (N Y)* **11**, 1256-1261.
19. Chockalingam K & Zhao H (2005) Creating new specific ligand-receptor pairs for transgene regulation. *Trends Biotechnol* **23**, 333-335.
20. Chockalingam K, Chen Z, Katzenellenbogen JA & Zhao H (2005) Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches. *Proc Natl Acad Sci U S A* **102**, 5691-5696.
21. Panguluri SK, Kumar P & Palli SR (2006) Functional characterization of ecdysone receptor gene switches in mammalian cells. *FEBS J* **273**, 5550-5563.
22. Doyle DF, Braasch DA, Jackson LK, Weiss HE, Boehm MF, Mangelsdorf DJ & Corey DR (2001) Engineering orthogonal ligand-receptor pairs from "near drugs". *J Am Chem Soc* **123**, 11367-11371.
23. Peet DJ, Doyle DF, Corey DR & Mangelsdorf DJ (1998) Engineering novel specificities for ligand-activated transcription in the nuclear hormone receptor RXR. *Chem Biol* **5**, 13-21.
24. Schwimmer LJ, Rohatgi P, Azizi B, Seley KL & Doyle DF (2004) Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules. *Proc Natl Acad Sci U S A* **101**, 14707-14712.
25. Tedesco R, Thomas JA, Katzenellenbogen BS & Katzenellenbogen JA (2001) The estrogen receptor: a structure-based approach to the design of new specific hormone-receptor combinations. *Chem Biol* **8**, 277-287.
26. Hassan AQ & Koh JT (2006) A functionally orthogonal ligand-receptor pair created by targeting the allosteric mechanism of the thyroid hormone receptor. *J Am Chem Soc* **128**, 8868-8874.
27. Koh JT (2002) Engineering selectivity and discrimination into ligand-receptor interfaces. *Chem Biol* **9**, 17-23.
28. Koh JT PM, Tomic-Canic M, McDaniel C (1999) Selective Regulation of Gene Expression Using Rationally-Modified Retinoic Acid Receptors. *Journal of the American Chemical Society* **121**, 1984-1985.
29. Shi Y & Koh JT (2002) Functionally orthogonal ligand-receptor pairs for the selective regulation of gene expression generated by manipulation of charged residues at the ligand-receptor interface of ER alpha and ER beta. *J Am Chem Soc* **124**, 6921-6928.

30. McInerney EM, Tsai MJ, O'Malley BW & Katzenellenbogen BS (1996) Analysis of estrogen receptor transcriptional enhancement by a nuclear hormone receptor coactivator. *Proc Natl Acad Sci U S A* **93**, 10069-10073.
31. Benecke A, Chambon P & Gronemeyer H (2000) Synergy between estrogen receptor alpha activation functions AF1 and AF2 mediated by transcription intermediary factor TIF2. *EMBO Rep* **1**, 151-157.
32. Webb P, Nguyen P, Shinsako J, Anderson C, Feng W, Nguyen MP, Chen D, Huang SM, Subramanian S, McKinerney E, Katzenellenbogen BS, Stallcup MR & Kushner PJ (1998) Estrogen receptor activation function 1 works by binding p160 coactivator proteins. *Mol Endocrinol* **12**, 1605-1618.
33. Kobayashi Y, Kitamoto T, Masuhiro Y, Watanabe M, Kase T, Metzger D, Yanagisawa J & Kato S (2000) p300 mediates functional synergism between AF-1 and AF-2 of estrogen receptor alpha and beta by interacting directly with the N-terminal A/B domains. *J Biol Chem* **275**, 15645-15651.
34. Onate SA, Boonyaratanakornkit V, Spencer TE, Tsai SY, Tsai MJ, Edwards DP & O'Malley BW (1998) The steroid receptor coactivator-1 contains multiple receptor interacting and activation domains that cooperatively enhance the activation function 1 (AF1) and AF2 domains of steroid receptors. *J Biol Chem* **273**, 12101-12108.
35. Bommer M, Benecke A, Gronemeyer H & Rochette-Egly C (2002) TIF2 mediates the synergy between RARalpha 1 activation functions AF-1 and AF-2. *J Biol Chem* **277**, 37961-37966.
36. Tian H, Mahajan MA, Wong CT, Habeos I & Samuels HH (2006) The N-Terminal A/B domain of the thyroid hormone receptor-beta2 isoform influences ligand-dependent recruitment of coactivators to the ligand-binding domain. *Mol Endocrinol* **20**, 2036-2051.
37. Germain P, Staels B, Dacquet C, Spedding M & Laudet V (2006) Overview of nomenclature of nuclear receptors. *Pharmacol Rev* **58**, 685-704.
38. Kumar V, Green S, Staub A & Chambon P (1986) Localisation of the oestradiol-binding and putative DNA-binding domains of the human oestrogen receptor. *EMBO J* **5**, 2231-2236.
39. Luisi BF, Xu WX, Otwinowski Z, Freedman LP, Yamamoto KR & Sigler PB (1991) Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature* **352**, 497-505.
40. Schwabe JW, Chapman L, Finch JT & Rhodes D (1993) The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements. *Cell* **75**, 567-578.

41. Umesono K & Evans RM (1989) Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell* **57**, 1139-1146.
42. Black BE, Holaska JM, Rastinejad F & Paschal BM (2001) DNA binding domains in diverse nuclear receptors function as nuclear export signals. *Curr Biol* **11**, 1749-1758.
43. Warnmark A, Treuter E, Wright AP & Gustafsson JA (2003) Activation functions 1 and 2 of nuclear receptors: molecular strategies for transcriptional activation. *Mol Endocrinol* **17**, 1901-1909.
44. Moras D & Gronemeyer H (1998) The nuclear receptor ligand-binding domain: structure and function. *Curr Opin Cell Biol* **10**, 384-391.
45. Germain P, Kammerer S, Perez E, Peluso-Iltis C, Tortolani D, Zusi FC, Starrett J, Lapointe P, Daris JP, Marinier A, de Lera AR, Rochel N & Gronemeyer H (2004) Rational design of RAR-selective ligands revealed by RARbeta crystal structure. *EMBO Rep* **5**, 877-882.
46. Charles JP, Shinoda T & Chinzei Y (1999) Characterization and DNA-binding properties of GRF, a novel monomeric binding orphan receptor related to GCNF and betaFTZ-F1. *Eur J Biochem* **266**, 181-190.
47. Rastinejad F, Wagner T, Zhao Q & Khorasanizadeh S (2000) Structure of the RXR-RAR DNA-binding complex on the retinoic acid response element DR1. *EMBO J* **19**, 1045-1054.
48. Rastinejad F, Perlmann T, Evans RM & Sigler PB (1995) Structural determinants of nuclear receptor assembly on DNA direct repeats. *Nature* **375**, 203-211.
49. Shaffer PL & Gewirth DT (2002) Structural basis of VDR-DNA interactions on direct repeat response elements. *EMBO J* **21**, 2242-2252.
50. Zhao Q, Chasse SA, Devarakonda S, Sierk ML, Ahvazi B & Rastinejad F (2000) Structural basis of RXR-DNA interactions. *J Mol Biol* **296**, 509-520.
51. Pratt WB & Toft DO (1997) Steroid receptor interactions with heat shock protein and immunophilin chaperones. *Endocr Rev* **18**, 306-360.
52. Rosenfeld MG, Lunyak VV & Glass CK (2006) Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. *Genes Dev* **20**, 1405-1428.
53. Perissi V & Rosenfeld MG (2005) Controlling nuclear receptors: the circular logic of cofactor cycles. *Nat Rev Mol Cell Biol* **6**, 542-554.
54. Privalsky ML (2004) The role of corepressors in transcriptional regulation by nuclear hormone receptors. *Annu Rev Physiol* **66**, 315-360.

55. Guenther MG, Barak O & Lazar MA (2001) The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3. *Mol Cell Biol* **21**, 6091-6101.
56. Love JD, Gooch JT, Nagy L, Chatterjee VK & Schwabe JW (2000) Transcriptional repression by nuclear receptors: mechanisms and role in disease. *Biochem Soc Trans* **28**, 390-396.
57. Nagy L, Kao HY, Chakravarti D, Lin RJ, Hassig CA, Ayer DE, Schreiber SL & Evans RM (1997) Nuclear receptor repression mediated by a complex containing SMRT, mSin3A, and histone deacetylase. *Cell* **89**, 373-380.
58. Nagy L, Kao HY, Love JD, Li C, Banayo E, Gooch JT, Krishna V, Chatterjee K, Evans RM & Schwabe JW (1999) Mechanism of corepressor binding and release from nuclear hormone receptors. *Genes Dev* **13**, 3209-3216.
59. Blondel A, Renaud JP, Fischer S, Moras D & Karplus M (1999) Retinoic acid receptor: a simulation analysis of retinoic acid binding and the resulting conformational changes. *J Mol Biol* **291**, 101-115.
60. Weatherman RV, Fletterick RJ & Scanlan TS (1999) Nuclear-receptor ligands and ligand-binding domains. *Annu Rev Biochem* **68**, 559-581.
61. Egea PF, Klaholz BP & Moras D (2000) Ligand-protein interactions in nuclear receptors of hormones. *FEBS Lett* **476**, 62-67.
62. Nagy L & Schwabe JW (2004) Mechanism of the nuclear receptor molecular switch. *Trends Biochem Sci* **29**, 317-324.
63. Chen D, Huang SM & Stallcup MR (2000) Synergistic, p160 coactivator-dependent enhancement of estrogen receptor function by CARM1 and p300. *J Biol Chem* **275**, 40810-40816.
64. McKenna NJ & O'Malley BW (2002) Combinatorial control of gene expression by nuclear receptors and coregulators. *Cell* **108**, 465-474.
65. Vo N & Goodman RH (2001) CREB-binding protein and p300 in transcriptional regulation. *J Biol Chem* **276**, 13505-13508.
66. Darimont BD, Wagner RL, Apriletti JW, Stallcup MR, Kushner PJ, Baxter JD, Fletterick RJ & Yamamoto KR (1998) Structure and specificity of nuclear receptor-coactivator interactions. *Genes Dev* **12**, 3343-3356.
67. Heery DM, Kalkhoven E, Hoare S & Parker MG (1997) A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **387**, 733-736.
68. Gurevich I, Flores AM & Aneskievich BJ (2007) Corepressors of agonist-bound nuclear receptors. *Toxicol Appl Pharmacol* **223**, 288-298.

69. Brzozowski AM, Pike AC, Dauter Z, Hubbard RE, Bonn T, Engstrom O, Ohman L, Greene GL, Gustafsson JA & Carlquist M (1997) Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* **389**, 753-758.
70. Leonhardt SA & Edwards DP (2002) Mechanism of action of progesterone antagonists. *Exp Biol Med (Maywood)* **227**, 969-980.
71. Hodgson MC, Shen HC, Hollenberg AN & Balk SP (2008) Structural basis for nuclear receptor corepressor recruitment by antagonist-liganded androgen receptor. *Mol Cancer Ther* **7**, 3187-3194.
72. Xu HE, Stanley TB, Montana VG, Lambert MH, Shearer BG, Cobb JE, McKee DD, Galardi CM, Plunket KD, Nolte RT, Parks DJ, Moore JT, Kliewer SA, Willson TM & Stimmel JB (2002) Structural basis for antagonist-mediated recruitment of nuclear co-repressors by PPARalpha. *Nature* **415**, 813-817.
73. Oberfield JL, Collins JL, Holmes CP, Goreham DM, Cooper JP, Cobb JE, Lenhard JM, Hull-Ryde EA, Mohr CP, Blanchard SG, Parks DJ, Moore LB, Lehmann JM, Plunket K, Miller AB, Milburn MV, Kliewer SA & Willson TM (1999) A peroxisome proliferator-activated receptor gamma ligand inhibits adipocyte differentiation. *Proc Natl Acad Sci U S A* **96**, 6102-6106.
74. Jackson TA, Richer JK, Bain DL, Takimoto GS, Tung L & Horwitz KB (1997) The partial agonist activity of antagonist-occupied steroid receptors is controlled by a novel hinge domain-binding coactivator L7/SPA and the corepressors N-CoR or SMRT. *Mol Endocrinol* **11**, 693-705.
75. Deroo BJ & Korach KS (2006) Estrogen receptors and human disease. *J Clin Invest* **116**, 561-570.
76. Love RR, Mazess RB, Barden HS, Epstein S, Newcomb PA, Jordan VC, Carbone PP & DeMets DL (1992) Effects of tamoxifen on bone mineral density in postmenopausal women with breast cancer. *N Engl J Med* **326**, 852-856.
77. Ismail SM (1994) Effects of tamoxifen on uterus. *Lancet* **344**, 622-623.
78. Glass CK (1994) Differential recognition of target genes by nuclear receptor monomers, dimers, and heterodimers. *Endocr Rev* **15**, 391-407.
79. Laudet V, Hanni C, Coll J, Catzeflis F & Stehelin D (1992) Evolution of the nuclear receptor gene superfamily. *EMBO J* **11**, 1003-1013.
80. Leid M, Kastner P, Lyons R, Nakshatri H, Saunders M, Zacharewski T, Chen JY, Staub A, Garnier JM, Mader S & et al. (1992) Purification, cloning, and RXR identity of the HeLa cell factor with which RAR or TR heterodimerizes to bind target sequences efficiently. *Cell* **68**, 377-395.

81. Shulman AI & Mangelsdorf DJ (2005) Retinoid x receptor heterodimers in the metabolic syndrome. *N Engl J Med* **353**, 604-615.
82. Nagy L, Thomazy VA, Heyman RA & Davies PJ (1998) Retinoid-induced apoptosis in normal and neoplastic tissues. *Cell Death Differ* **5**, 11-19.
83. Mangelsdorf DJ, Borgmeyer U, Heyman RA, Zhou JY, Ong ES, Oro AE, Kakizuka A & Evans RM (1992) Characterization of three RXR genes that mediate the action of 9-cis retinoic acid. *Genes Dev* **6**, 329-344.
84. Mangelsdorf DJ, Ong ES, Dyck JA & Evans RM (1990) Nuclear receptor that identifies a novel retinoic acid response pathway. *Nature* **345**, 224-229.
85. Dolle P, Fraulob V, Kastner P & Chambon P (1994) Developmental expression of murine retinoid X receptor (RXR) genes. *Mech Dev* **45**, 91-104.
86. Hamada K, Gleason SL, Levi BZ, Hirschfeld S, Appella E & Ozato K (1989) H-2RIIBP, a member of the nuclear hormone receptor superfamily that binds to both the regulatory element of major histocompatibility class I genes and the estrogen response element. *Proc Natl Acad Sci U S A* **86**, 8289-8293.
87. Yu VC, Delsert C, Andersen B, Holloway JM, Devary OV, Naar AM, Kim SY, Boutin JM, Glass CK & Rosenfeld MG (1991) RXR beta: a coregulator that enhances binding of retinoic acid, thyroid hormone, and vitamin D receptors to their cognate response elements. *Cell* **67**, 1251-1266.
88. Haugen BR, Brown NS, Wood WM, Gordon DF & Ridgway EC (1997) The thyrotrope-restricted isoform of the retinoid-X receptor-gamma1 mediates 9-cis-retinoic acid suppression of thyrotropin-beta promoter activity. *Mol Endocrinol* **11**, 481-489.
89. Chiang MY, Misner D, Kempermann G, Schikorski T, Giguere V, Sucov HM, Gage FH, Stevens CF & Evans RM (1998) An essential role for retinoid receptors RARbeta and RXRgamma in long-term potentiation and depression. *Neuron* **21**, 1353-1361.
90. Mader S, Chen JY, Chen Z, White J, Chambon P & Gronemeyer H (1993) The patterns of binding of RAR, RXR and TR homo- and heterodimers to direct repeats are dictated by the binding specificities of the DNA binding domains. *EMBO J* **12**, 5029-5041.
91. Mangelsdorf DJ, Umesono K, Kliewer SA, Borgmeyer U, Ong ES & Evans RM (1991) A direct repeat in the cellular retinol-binding protein type II gene confers differential regulation by RXR and RAR. *Cell* **66**, 555-561.
92. Shulman AI, Larson C, Mangelsdorf DJ & Ranganathan R (2004) Structural determinants of allosteric ligand activation in RXR heterodimers. *Cell* **116**, 417-429.

93. Westin S, Kurokawa R, Nolte RT, Wisely GB, McInerney EM, Rose DW, Milburn MV, Rosenfeld MG & Glass CK (1998) Interactions controlling the assembly of nuclear-receptor heterodimers and co-activators. *Nature* **395**, 199-202.
94. Mertz JR, Shang E, Piantedosi R, Wei S, Wolgemuth DJ & Blaner WS (1997) Identification and characterization of a stereospecific human enzyme that catalyzes 9-cis-retinol oxidation. A possible role in 9-cis-retinoic acid formation. *J Biol Chem* **272**, 11744-11749.
95. Romert A, Tuvendal P, Simon A, Dencker L & Eriksson U (1998) The identification of a 9-cis retinol dehydrogenase in the mouse embryo reveals a pathway for synthesis of 9-cis retinoic acid. *Proc Natl Acad Sci U S A* **95**, 4404-4409.
96. de Urquiza AM, Liu S, Sjoberg M, Zetterstrom RH, Griffiths W, Sjovall J & Perlmann T (2000) Docosahexaenoic acid, a ligand for the retinoid X receptor in mouse brain. *Science* **290**, 2140-2144.
97. Lampen A, Meyer S & Nau H (2001) Phytanic acid and docosahexaenoic acid increase the metabolism of all-trans-retinoic acid and CYP26 gene expression in intestinal cells. *Biochim Biophys Acta* **1521**, 97-106.
98. Lemotte PK, Keidel S & Apfel CM (1996) Phytanic acid is a retinoid X receptor ligand. *Eur J Biochem* **236**, 328-333.
99. Renaud JP, Rochel N, Ruff M, Vivat V, Chambon P, Gronemeyer H & Moras D (1995) Crystal structure of the RAR-gamma ligand-binding domain bound to all-trans retinoic acid. *Nature* **378**, 681-689.
100. Bourguet W, Vivat V, Wurtz JM, Chambon P, Gronemeyer H & Moras D (2000) Crystal structure of a heterodimeric complex of RAR and RXR ligand-binding domains. *Mol Cell* **5**, 289-298.
101. Egea PF, Mitschler A, Rochel N, Ruff M, Chambon P & Moras D (2000) Crystal structure of the human RXRalpha ligand-binding domain bound to its natural ligand: 9-cis retinoic acid. *Embo J* **19**, 2592-2601.

CHAPTER 2

THE POWER OF GENETIC SELECTION

2.1 Genetic Selection

Genetic selection is one of the most powerful tools used to analyze proteins with novel functions, such as engineered NR responses to non-natural ligands. In genetic selection, the survival of a host cell is linked to a desired entity, such as a desired macromolecule's function [1-7]. The use of genetic selection in high-throughput assays has been shown to greatly benefit drug discovery, protein engineering, enzymology, and proteomics [8-10]. The use of high-throughput assays coupled with genetic selection to evolve proteins with new properties has also provided a powerful tool for understanding both protein structure and function [1-7].

Analyzing variants of large combinatorial protein libraries can be done using two methods, genetic screening or genetic selection. When using genetic screening techniques, one must evaluate all variants of the library to determine which variants have the desired function. In facilitated screening, functional variants will have a distinctive phenotype, however all members of the library, functional and non-functional, will always be present [11]. When using genetic selection techniques, since the survival of the host cell is linked to a particular function, only functional variants survive. This eliminates the need to evaluate all variants, functional and non-functional, allowing for easier analysis of large protein libraries. Variants can be analyzed faster in genetic selection systems, due to the elimination of evaluating uninteresting variants, allowing for the analysis of much larger library sizes. Using the best screening protocol in

Escherichia coli (*E. coli*), only 10^5 library members can be assayed, however, with selection, up to 10^{10} variants can be assayed in *E. coli* [12].

One of the problems of genetic selection systems is that an appropriate selection strategy, choice of host and complementation strategies, must be developed for every application [13]. Thus, creating a generalizable selection method for various applications can pose to be a challenge for genetic selection. However, with the use of various microbial strains and the development of more and more selection techniques many genetic selection systems have been developed [3, 7-9, 11, 12, 14-22]. Commonly in genetic selection systems, cell survival is linked to enzyme function. In principle, any enzyme activity can be selected for, provided cell survival can be linked to the catalysis of the desired reaction. One strategy for creating a selection technique is the use of a metabolic requirement for a desired enzyme activity. One such strategy is the use of chorismate mutase activity [23]. Briefly, chorismate mutase catalyzes the first step in the biosynthesis of phenylalanine and tyrosine [24]. To create a chorismate mutase based genetic selection system, the genes encoding the chorismate mutase were altered in *E. coli*, and as a result survival of the strain in media lacking phenylalanine and tyrosine requires an added source of the chorismate mutase. This is usually accomplished by the transformation of a plasmid carrying the gene encoding this enzyme. In selection systems containing a functional chorismate mutase, cells are able to survive in media lacking phenylalanine and tyrosine.

Many genetic selection systems have been developed to evaluate numerous macromolecular interactions, ranging from protein-DNA or RNA, protein-protein, and protein-ligand interactions [9, 15, 18-20]. A popular genetic selection technique to

evaluate such interactions is the use to two and three hybrid systems, which will be discussed in the remainder of this chapter.

2.2 Two Hybrid Systems

With the advancements of genetic selection as one of the most powerful tools to both discover and analyze protein function, two-hybrid systems (particularly the yeast two-hybrid system, Y2H) have emerged as one of the most successful genetic selection/screening based methods used to evaluate protein function. Protein-protein interactions play a crucial role in virtually every cellular process, including DNA replication, transcription and translation, as well as metabolism. Additionally the alterations of these interactions are known to contribute to a variety of disease. With the advancements in the human genome project, the idea that understanding protein function requires an understanding of protein-protein interactions has lead researchers to look for proteins which interact with their proteins of interest. The Y2H provides an efficient system to analyze such interactions.

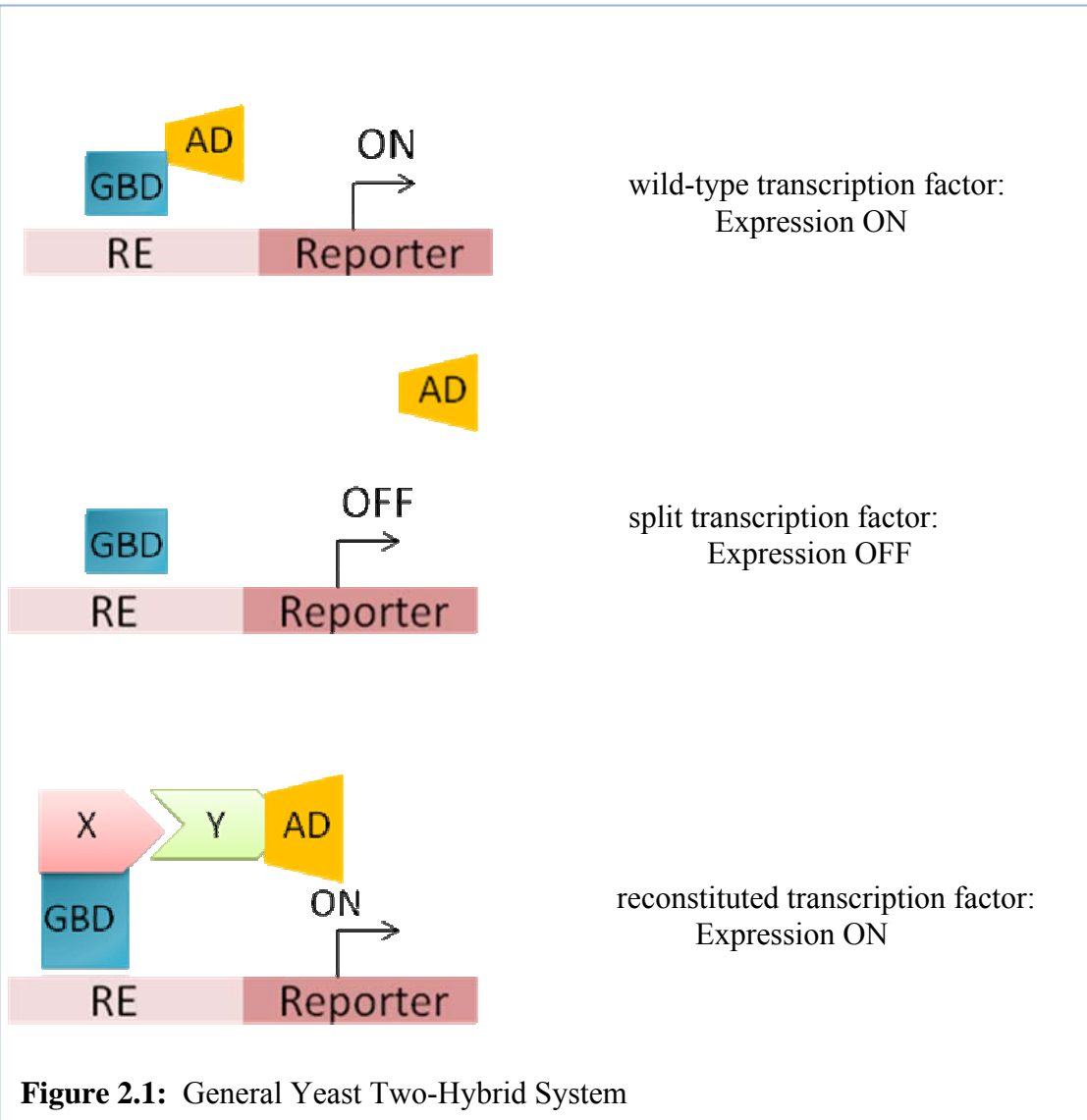
The basic idea of the two-hybrid method is to control the expression of a gene by splitting a protein into two halves that do not work independently but will work when brought together again, most commonly a transcription factor protein (TF) because it is capable of controlling the activation and/or repression of a particular gene of interest when functional. The TF is generally split into two domains, a DNA binding domain (DBD) and an activation domain (AD). When the two domains are expressed as fusion proteins, or “hybrids”, with two other proteins with a sufficient affinity for each other, the function of the TF is restored. The protein fused to the DBD of the TF is referred to as the “bait” (X). The protein fused to the AD of the TF is the “prey” (Y). If the bait and prey

proteins are able to interact, bringing together the DBD and AD of the TF, a functional TF will be reconstituted restoring the function of the TF (Figure 2.1) [25].

Genetic screening or selection is used to detect the reconstitution of the TF by transcriptional activation of a reporter gene. The reporters usually generate a colorimetric or fluorescent readout (screening) or allows growth in selective media (selection). Two common examples of reporter genes include the *LacZ* and *HIS3* genes. The *LacZ* gene is used as a screen and encodes β -galactosidase, an enzyme that hydrolyzes ortho-nitrophenyl- β -galactoside (ONPG) into ortho-nitrophenol (ONP), a yellow product. The *HIS3* gene is a selection marker which encodes for the enzyme imidazole-glycerolphosphate dehydratase. This enzyme is crucial to histidine biosynthesis and allows cells lacking a functional *HIS3* gene to grow in media lacking histidine.

The first two-hybrid system was developed in yeast by Fields and Song almost two decades ago as a “proof-of-concept” that transcriptional readout could be used as a tool to investigate protein-protein interactions [26]. They used the yeast TF, GAL4, and fused its DBD to a serine-threonine-specific kinase, SNF1, and the GAL4 AD was fused to SNF4, a protein previously shown to interact with SNF1 and required for its maximal activity [27]. When both hybrid proteins were expressed in a yeast strain with the genomic GAL4 gene deleted, they were able to obtain transcriptional activation of a *lacZ* gene placed downstream of a GAL4 promoter. Since this development, thousands of publications have used this system to characterize protein-protein interactions (reviewed in [28-31]).

Generally, to characterize protein-protein interactions with two-hybrid systems,



proteins of interest are fused to the DBD of GAL4. This fusion is then used as “bait” to screen, or select, through libraries of “prey” proteins fused to the AD of GAL4. If a variant from the library is able to interact with the bait protein, the DBD and AD will be brought together resulting in transcriptional activation of a reporter gene. Since the development of the Y2H by Fields, numerous improvements have been integrated which have increased its applicability. With these improvements the Y2H has been shown not only to determine protein-protein interactions [32-35], but has also been used to analyze protein function, decipher protein networks, and to detect other protein-macromolecular interactions such as small molecules and RNA [8, 14, 15, 18, 34-37]. The remainder of this chapter will focus on adaptations of the Fields Y2H in both yeast and bacteria.

Since the pioneering of the Y2H, this technique has been adapted, improved, and diversified. Improvements to the Y2H include vector refinement [38] and new strain developments with various selectable markers [39]. With its adaptations, this system has also gone on to be used for a variety of applications including genome mapping and drug discovery. The reverse two-hybrid system (rY2H) and the yeast three-hybrid systems (Y3H) are two of the major Y2H variants used to probe protein-small molecule interactions essential for drug discovery.

In addition to identifying protein-protein interactions it is also crucial to understand the regulation of protein-protein interactions as well as characterize structure-function relationships. To do this, events which dissociate protein-protein interactions need to be investigated. Structure-function relations can be investigated by detecting mutations in either protein partner which will cause dissociation. Regulatory

mechanisms can be investigated using small molecules to mediate dissociation of the protein partners. Classical Y2H that employ positive selection does not account for these occurrences. rY2H are able to detect mutations or small molecules that disrupt protein-protein interactions. Reverse systems make use of yeast strains engineered for negative selection. In negative selection, expression of interacting hybrid proteins results in the expression of a counterselectable marker which is toxic to the cells under specific conditions. Under these conditions disruption of the protein-protein interactions, by mutations or small molecules, provides an advantage, generally cell viability.

The Y3H extends the role of traditional Y2H to not only investigate protein-protein interactions, but to detect protein-RNA and protein-small molecule interactions as well [37, 40, 41]. Whereas the rY2H assess the disruption of protein-protein interactions, Y3H allow researchers to screen, or select, for proteins which directly interact with other macromolecules such as RNA or a specific small molecule *in vivo* [8, 37, 40-44]. In these systems, the RNA or small molecules serve as a bridge between the two interacting proteins.

The Y3H has also become a popular tool for evaluate small molecules for drug discovery [37, 40]. Additionally, the use of small molecules to regulate gene expression has various clinical applications, specifically gene therapy. The ability to use a small molecule to control gene expression in a dose dependent manner has even more implications and is an extremely valuable tool. Multiple Y3H assays have been developed for evaluating small molecule-protein, specifically NRs, interactions [8, 15, 36, 37, 45-52]. These assays have not only created tools for creating and developing

small molecule based regulators of gene expression, but have also become useful tools for enzyme and protein engineering as well [8, 45, 48, 51].

Y3Hs developed for protein-small molecule interactions, use the small molecule ligand as a bridge, or linker, between the bait and prey proteins. In most applications, binding the small molecule causes the two proteins to dimerize, resulting in transcriptional activation. These small molecules are referred to as chemical inducers of dimerization (CID). Cornish and colleagues developed aY3H which uses a genetic screen to link a heterodimeric small molecule to enzyme catalysis [53]. The Cornish Y3H consists of a heterodimeric small molecule which bridges a DBD-receptor fusion protein and an AD-receptor fusion protein. This approach is generalizable because it can be adapted to essentially any enzyme simply by changing the substrate linking the bait and prey proteins. The Cornish group has been successful in using this Y3H to screen through libraries of CIDs as a means to characterize enzymes [53, 54]

2.3 Chemical Complementation

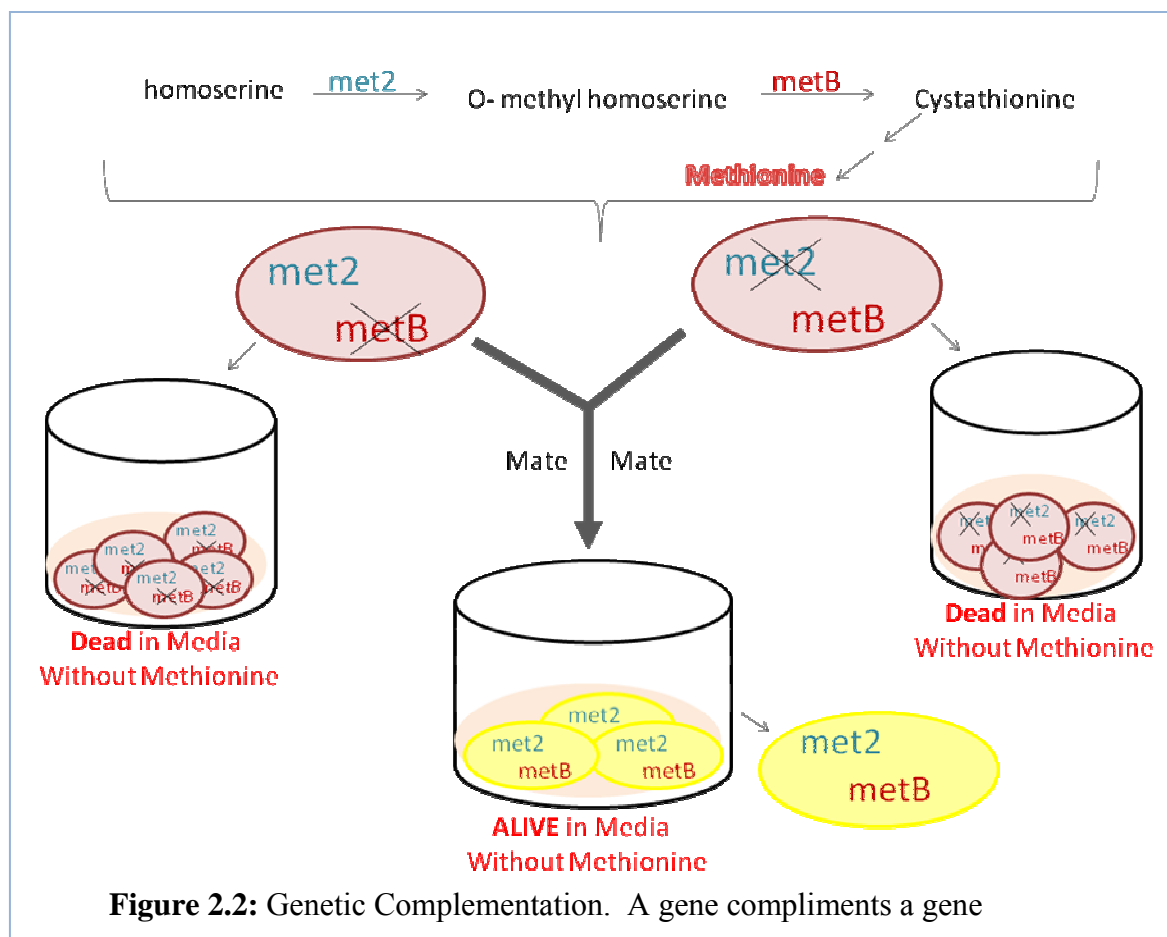
Creating a generalizable high-throughput assay for the detection of protein variants with improved or novel activity is a tool that would greatly benefit protein engineering. As discussed above, genetic screening as well as selection can be used to facilitate such searches for both proteins and/or ligands from these libraries. Successful isolation of desirable proteins or ligands from these libraries depends both on the quality and diversity of the library as well as the power of the screen or the selection. However, it is important to note that in a screening assay one must evaluate all variants, both functional and non-functional, whereas a selection assay links function of a protein to cell viability and only functional variants survive, eliminating the need to evaluate non-

functional variants. There are a lack of generalizable selections systems, thus to address this issue we have developed chemical complementation (CC), a system modeled after the Y2H, which uses genetic selection to link the survival of a cell to the presence of a specific small molecule in *S. cerevisiae* [14, 44].

CC can be compared to classic genetic complementation in that both systems use complementation, either by a gene or small molecule, to link genetic selection to survival of a yeast cell. In classic genetic complementation, genetic selection is used to link the survival of the yeast cell to a functional gene that complements a non-functional gene (e.g. an auxotroph). However, with CC genetic selection is used to link the survival of the yeast cell to the presence of a small molecule (chemical) through binding a nuclear receptor and activating transcription of a functional gene (Figures 2.2 and 2.3).

CC is a three component system that comprises a human NR protein-GAL4 DBD fusion protein, a NR coactivator protein-GAL4 AD fusion protein, and a small molecule ligand. Upon binding of its specific small molecule ligand, the nuclear receptor undergoes a conformational change recruiting the coactivator and activating transcription of a selectable marker (Figure 2.3). These hybrid proteins are expressed in the *S. cerevisiae* strain PJ69-4A. PJ69-4A is a third generation Y2H strain which contains separate GAL4 response elements (GAL4 REs) upstream of the genetic selection and screening genes, *HIS3*, *ADE2*, and *LacZ* [39]. The system was developed using various nuclear receptors-ligand pairs, to include RXR-9cRA, the pregnane X receptor (PXR) with paclitaxel, and the liver X receptor (LXR) with oxysterols.

Because NR ligands can act as both agonists and antagonists, CC was also developed to select for antagonist ligands as well using negative selection [44]. NR



agonists are small molecules which bind, cause a conformational change, and activate transcription. Antagonists are small molecules which bind but prevent activation. Potent antagonists are able to displace agonist, resulting in an inactive NR conformation. CC uses positive selection to select for NR agonist. In negative chemical complementation (NCC), negative selection is used to select for antagonists. To develop NCC, a new *S. cerevisiae* strain had to be constructed, BAPJ69, in which a negative selection marker (*URA3*) was put under the control of a GAL4 RE. The *URA3* gene encodes for orotidine-5'-phosphate decarboxylase, an enzyme required for uracil biosynthesis and allows cells to survive in media lacking uracil. However, orotidine-5'-phosphate decarboxylase also catalyzes the conversion of a non-toxic 5-fluoroorotic acid (FOA) into toxic 5-fluorouracil [55]. With functional NR-agonist pairs, cells will die in media lacking uracil and containing FOA. But, if a functional NR-antagonist pair is present, orotidine-5'-phosphate decarboxylase will not be expressed, thus non-toxic 5-fluoroorotic acid (FOA) cannot be converted into toxic 5-fluorouracil, allowing cells to live under these conditions (Figure 2.4).

The CC system developed in yeast is highly sensitive (producing growth at 10 nM ligand concentration), has zero background (no growth without ligand) and high dynamic range (growth density and time equivalent to Gal4 induced growth). This system offers a general method of engineering receptors that activate transcription in response to arbitrary small molecules. We have shown that yeast has proven to be a very efficient method for analyzing libraries for functional ligand-receptor pairs [8]. Schwimmer et al. utilized CC to screen libraries of mutant RXR variants to bind the synthetic ligand, LG335. RXR wild-type (wt) binds its natural ligand, 9cRA, with an EC_{50} of 500 nM

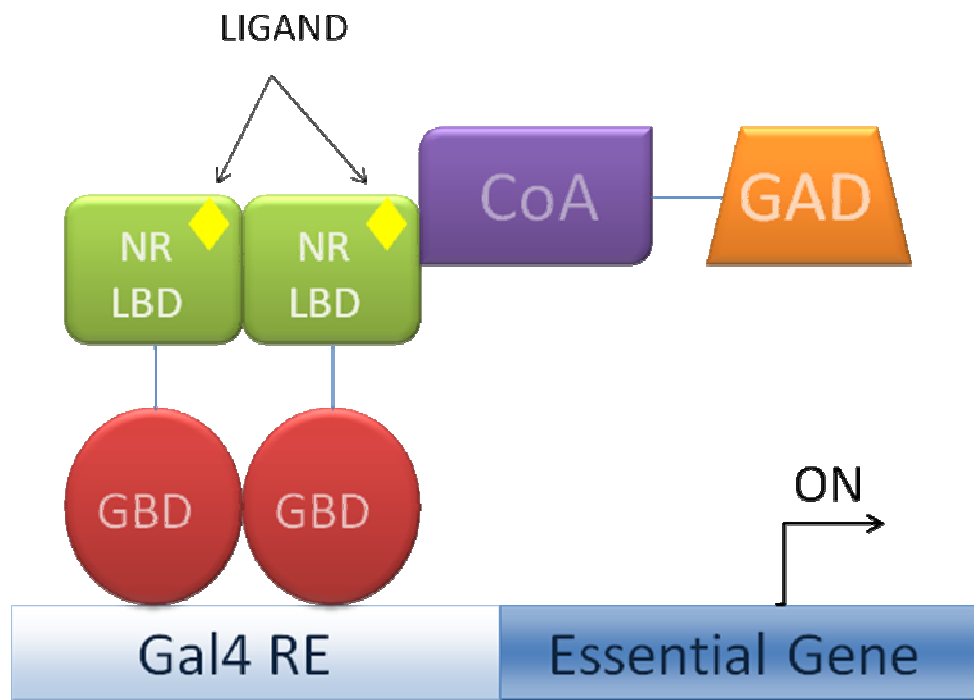


Figure 2.3: Chemical Complementation

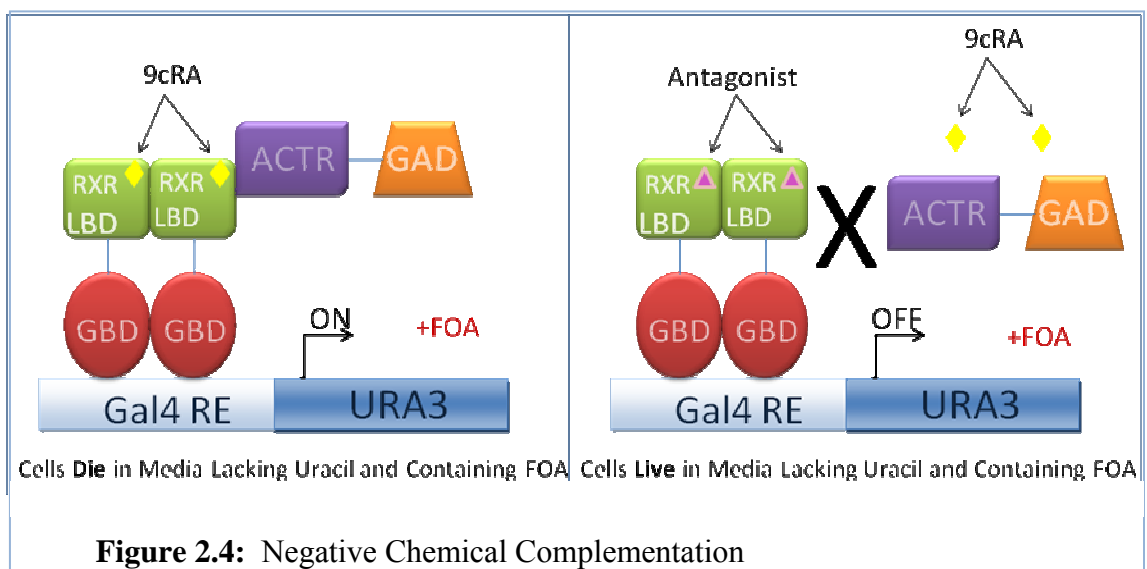


Figure 2.4: Negative Chemical Complementation

and the synthetic ligand, LG335, with an EC_{50} greater than $10\ \mu\text{M}$. Schwimmer et al. was able to successfully use CC to select through a library of approximately 400,000 RXR variants to find the RXR variant, I268V;A272V;I310L;F313M, capable of binding and activating transcription with the synthetic ligand, LG335, with an EC_{50} of $40\ \text{nm}$ and activating transcription with the wt ligand, 9cRA, with an EC_{50} greater than $10\ \mu\text{M}$. CC is a powerful tool for analyzing NR-ligand interactions. This system can be extended to a variety of applications to include drug discovery, protein/enzyme engineering, and deciphering and assembling biosynthetic pathways.

2.4 Bacterial Two-Hybrid Systems

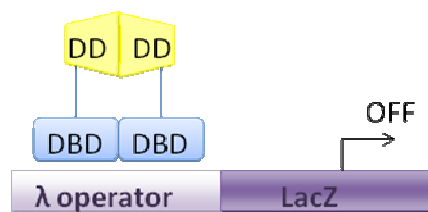
The Y2H systems has also been expanded and adapted to other organisms and hosts, particularly *E. coli*. Before the development of the bacterial two-hybrid systems (B2H), protein-macromolecule interactions were studied by mainly using the Y2H and bacterial-based phage display methods [56-60]. The development of a B2H to analyze protein-macromolecule interactions has advantages over Y2H and phage-display methods. First, B2H methods can facilitate rapid analysis of larger libraries than yeast, due to *E. coli*'s faster growth rate and higher transformation efficiency. B2H can also be faster than phage-display methods and allow for the investigation of larger proteins not readily displayed on phage surfaces.

B2H are either developed based on bacterial repressors, such as the bacteriophage lambda repressor (λCI), or transcriptional activation using RNA polymerase (RNAP). Repressor based systems typically use the LexA repressor or the λCI [61-64]. The λCI -B2H was first developed by Hu and colleagues. When bound to the λ operator, the λCI prevents expression of genes involved in the bacterial lytic program [62, 63]. This

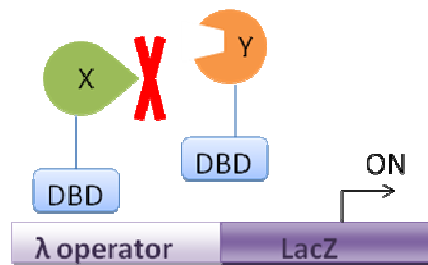
repressor is only active in the dimer form and each monomer contains a DBD and a dimerization domain (DD) [65]. Hence, the DBD alone is unable to function as a repressor due to its inability to dimerize. λ CI-B2H was constructed similarly to the Y2H by fusing the bait protein to a λ CI-DBD and the prey protein to another λ CI-DBD. Only systems with a functional bait-prey interaction are able to form a functional λ CI-DBD dimer, allowing repression of a *LacZ* gene downstream of λ promoter-operator [62, 63] (Figure 2.5). Resultantly, a decrease in ONPG to ONP conversion can be observed. Many groups have been able to use this system to analyze the dimerization of various proteins [63, 64, 66-68].

A major drawback of using the λ CI-B2H is that this system cannot be applied to screening for protein-protein interactions if the bait proteins are able to homodimerize. This situation would result in functional repression regardless of the presence of a prey-DBD hybrid. As a result of this, the LexA-B2H was developed. This system works similarly to the λ CI-B2H with the advantage that allows the system to overcome the limitation of not being able to assess bait proteins able to homodimerize [61]. In this system, instead of fusing the prey protein to the wild-type DBD of LexA, it is fused to a mutant form of the LexA DBD and controls a *LacZ* gene upstream of operator made up of half-sites for both the wild-type (wt) and mutant LexA DBDs. Resultantly, only heterodimers of the wt-DBD and the mutant-DBD are able to repress expression of *LacZ*.

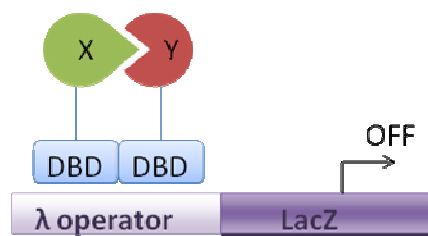
In contrast to repressor based B2H, there have also been a number of systems developed based on transcriptional activation using bacterial RNAP. This approach was first developed by Hochschild and coworkers, after discovering that interactions between



Active Repressor: Binds DNA as a dimer, blocking transcription



Inactive Repressor: DBDs unable to dimerize or bind DNA, allowing transcription to occur



Reconstituted Repressor: Able to dimerize and bind DNA blocking transcription

Figure 2.5: Repression Based Bacterial Two-Hybrid System (B2H)

DNA-bound activators and the C-terminal domain of the α -subunit of RNAP resulted in the N-terminal domain of the α -subunit of RNAP being able to recruit the RNAP core enzyme and initiate transcription of numerous genes [62, 69]. To develop this system the DBD was fused to a “bait” protein while the “prey” protein was fused to the N-terminal domain of the α -subunit (Figure 2.6). Thus, if the bait and prey proteins are able to interact, the *LacZ* gene, located downstream of DBD binding site and lac promoter construct, will be activated.

Joung and co-workers were able to develop a genetic selection based B2H for studying both protein-protein and protein-DNA interactions [12]. The Joung system was developed based off of the Hochschild genetic screen, where the *LacZ* marker was replaced with a *HIS3* selection marker. The *HIS3* gene is a selection marker which encodes for the enzyme imidazole-glycerolphosphate dehydratase. This enzyme is crucial to histidine biosynthesis and allows cells lacking a functional *HIS3* gene to grow in media lacking histidine. Only cells with a functional protein-protein interaction will survive in media lacking histidine, thus allowing for the evaluation of libraries $\sim 10^{10}$ in size as opposed to between 10^5 - 10^6 in size as observed with screens. Additionally, the Joung B2H is able to analyze protein-DNA interactions as well as protein-protein interactions. In this system, instead of the standard two hybrid system, in which the “prey” proteins are assayed for their ability to interact with a “bait” protein, Joung and colleagues used “bait” and “prey” proteins previously known to interact, the Gal4 and Gal11p proteins, and created a system capable of selecting through large randomized libraries of zinc finger DNA-binding domain variants for their ability to bind tightly and specifically to desired DNA binding sites.

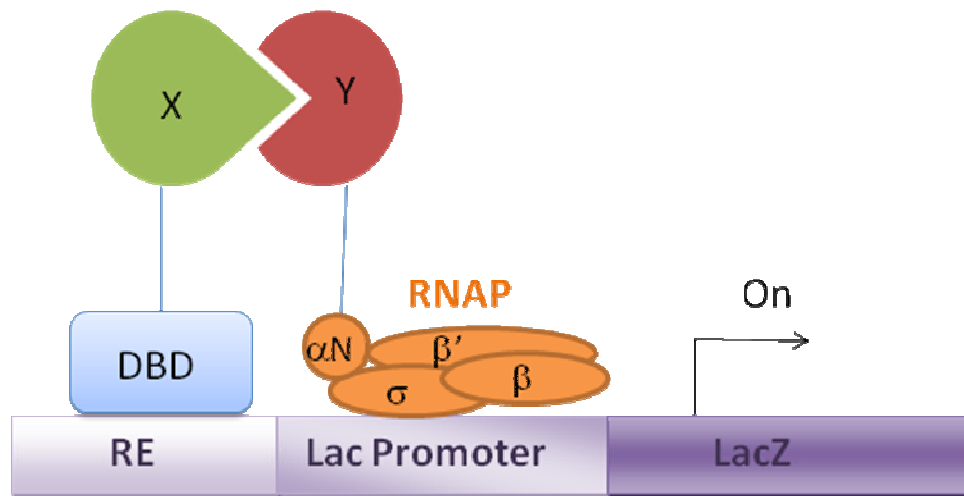


Figure 2.6: Activation Based Bacterial Two-Hybrid System



Figure 2.7: Wood's Bacterial One-Hybrid Construct

Although the B2H is not as well defined as the Y2H, the B2H can have numerous potential applications for in vivo analysis of protein-macromolecule interactions. An application currently developed in *E. coli* involving genetic selection and a unique version of the bacterial one-hybrid system (B1H) has been developed by Wood and coworkers. Wood and colleagues have developed a high-throughput assay to discover and characterize protein-small molecule interactions [21, 22, 70, 71]. In their optimized B1H, a single hybrid construct was developed which comprised the LBD of the estrogen receptor (ER) fused to a thymidylate synthase enzyme (TS), a stabilization domain (the inactive Mtu-RecA intein), and a solubilization domain (the maltose binding protein tag, MBP) (Figure 2.7). When this construct is expressed in a TS deficient *E. coli* strain at 34°C without a NR interacting small molecule, the cells are unable to grow in media lacking thymine. However, when an interacting small molecule is introduced into the same cells expressing a functional construct, cells are able to survive in media lacking thymine. Wood and colleagues hypothesize that this ligand induced growth is a result of the NRs conformational change induced by ligand binding. The fusion of the intein to the N terminus of the TS enzyme sterically blocks homodimerization, which is required for activation. The NR conformational changes cause a rearrangement of the intein structure, partially alleviating the blocked TS and enhancing TS function. Additionally, this system was able to evaluate both agonist and antagonist of ER, due to thymidylate synthases ability to function in both positive and negative selection. The drawback of this system is that it cannot be easily adapted to evaluate interactions other than protein-small molecule interactions.

2.5 Literature Cited

1. Dunham MJ, Badrane H, Ferea T, Adams J, Brown PO, Rosenzweig F & Botstein D (2002) Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* **99**, 16144-16149.
2. Gellerfors PL, Saltzgaber-Muller J & Douglas MG (1986) Selection by genetic complementation and characterization of the gene coding for the yeast porphobilinogen deaminase. *Biochem J* **240**, 673-677.
3. Lin H & Cornish VW (2002) Screening and selection methods for large-scale analysis of protein function. *Angew Chem Int Ed Engl* **41**, 4402-4425.
4. Miller AM, Savinelli EA, Couture SM, Hannigan GM, Han Z, Selden RF & Treco DA (1993) Recombination walking: genetic selection of clones from pooled libraries of yeast artificial chromosomes by homologous recombination. *Proc Natl Acad Sci U S A* **90**, 8118-8122.
5. O'Malley K & Douglas MG (1983) Selection and characterization of nuclear genes coding mitochondrial proteins: genetic complementation of yeast pet mutants. *Methods Enzymol* **97**, 344-355.
6. O'Malley K, Pratt P, Robertson J, Lilly M & Douglas MG (1982) Selection of the nuclear gene for the mitochondrial adenine nucleotide translocator by genetic complementation of the *op1* mutation in yeast. *J Biol Chem* **257**, 2097-2103.
7. Woycechowsky KJ & Hilvert D (2004) Deciphering enzymes. Genetic selection as a probe of structure and mechanism. *Eur J Biochem* **271**, 1630-1637.
8. Schwimmer LJ, Rohatgi P, Azizi B, Seley KL & Doyle DF (2004) Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules. *Proc Natl Acad Sci U S A* **101**, 14707-14712.
9. Norman TC, Smith DL, Sorger PK, Drees BL, O'Rourke SM, Hughes TR, Roberts CJ, Friend SH, Fields S & Murray AW (1999) Genetic selection of peptide inhibitors of biological pathways. *Science* **285**, 591-595.
10. Jackel C, Kast P & Hilvert D (2008) Protein design by directed evolution. *Annu Rev Biophys* **37**, 153-173.
11. Taylor SV, Kast P & Hilvert D (2001) Investigating and Engineering Enzymes by Genetic Selection. *Angew Chem Int Ed Engl* **40**, 3310-3335.
12. Joung JK, Ramm EI & Pabo CO (2000) A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. *Proc Natl Acad Sci U S A* **97**, 7382-7387.

13. Gildersleeve J, Janes J, Ulrich H, Yang P, Barbas C & Schultz PG (2002) Development of a genetic selection for catalytic antibodies. *Bioorg Med Chem Lett* **12**, 1691-1694.
14. Azizi B, Chang EI & Doyle DF (2003) Chemical complementation: small-molecule-based genetic selection in yeast. *Biochem Biophys Res Commun* **306**, 774-780.
15. Drees BL (1999) Progress and variations in two-hybrid and three-hybrid technologies. *Curr Opin Chem Biol* **3**, 64-70.
16. Katzenellenbogen JA & Katzenellenbogen BS (1996) Nuclear hormone receptors: ligand-activated regulators of transcription and diverse cell responses. *Chem Biol* **3**, 529-536.
17. Koh JT PM, Tomic-Canic M, McDaniel C (1999) Selective Regulation of Gene Expression Using Rationally-Modified Retinoic Acid Receptors. *Journal of the American Chemical Society* **121**, 1984-1985.
18. McNabb DS & Guarente L (1996) Genetic and biochemical probes for protein-protein interactions. *Curr Opin Biotechnol* **7**, 554-559.
19. Warbrick E (1997) Two's company, three's a crowd: the yeast two hybrid system for mapping molecular interactions. *Structure* **5**, 13-17.
20. Youderian P, Vershon A, Bouvier S, Sauer RT & Susskind MM (1983) Changing the DNA-binding specificity of a repressor. *Cell* **35**, 777-783.
21. Skretas G, Meligova AK, Villalonga-Barber C, Mitsiou DJ, Alexis MN, Michas-Screttas M, Steele BR, Screttas CG & Wood DW (2007) Engineered chimeric enzymes as tools for drug discovery: generating reliable bacterial screens for the detection, discovery, and assessment of estrogen receptor modulators. *J Am Chem Soc* **129**, 8443-8457.
22. Skretas G & Wood DW (2005) A bacterial biosensor of endocrine modulators. *J Mol Biol* **349**, 464-474.
23. Kast P, Asif-Ullah M, Jiang N & Hilvert D (1996) Exploring the active site of chorismate mutase by combinatorial mutagenesis and selection: the importance of electrostatic catalysis. *Proc Natl Acad Sci U S A* **93**, 5043-5048.
24. Haslam E (1993) *Shikimic acid : metabolism and metabolites* Wiley, Chichester ; New York ;.
25. Koegl M & Uetz P (2007) Improving yeast two-hybrid screening systems. *Brief Funct Genomic Proteomic* **6**, 302-312.
26. Fields S & Song O (1989) A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245-246.

27. Celenza JL & Carlson M (1986) A yeast gene that is essential for release from glucose repression encodes a protein kinase. *Science* **233**, 1175-1180.
28. Van Crieginge W & Beyaert R (1999) Yeast Two-Hybrid: State of the Art. *Biol Proced Online* **2**, 1-38.
29. McAlister-Henn L, Gibson N & Panisko E (1999) Applications of the yeast two-hybrid system. *Methods* **19**, 330-337.
30. Suter B, Kittanakom S & Stagljar I (2008) Two-hybrid technologies in proteomics research. *Curr Opin Biotechnol* **19**, 316-323.
31. Navaratnam DS (2009) Yeast two-hybrid screening to test for protein-protein interactions in the auditory system. *Methods Mol Biol* **493**, 257-268.
32. Serebriiskii IG, Fang R, Latypova E, Hopkins R, Vinson C, Joung JK & Golemis EA (2005) A combined yeast/bacteria two-hybrid system: development and evaluation. *Mol Cell Proteomics* **4**, 819-826.
33. Toby GG & Golemis EA (2001) Using the yeast interaction trap and other two-hybrid-based approaches to study protein-protein interactions. *Methods* **24**, 201-217.
34. Uetz P (2002) Two-hybrid arrays. *Curr Opin Chem Biol* **6**, 57-62.
35. Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P, Qureshi-Emili A, Li Y, Godwin B, Conover D, Kalbfleisch T, Vijayadamodar G, Yang M, Johnston M, Fields S & Rothberg JM (2000) A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* **403**, 623-627.
36. Griffith EC, Licitra EJ & Liu JO (2000) Yeast three-hybrid system for detecting ligand-receptor interactions. *Methods Enzymol* **328**, 89-103.
37. Licitra EJ & Liu JO (1996) A three-hybrid system for detecting small ligand-protein receptor interactions. *Proc Natl Acad Sci U S A* **93**, 12817-12821.
38. Roder KH, Wolf SS & Schweizer M (1996) Refinement of vectors for use in the yeast two-hybrid system. *Anal Biochem* **241**, 260-262.
39. James P, Halladay J & Craig EA (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144**, 1425-1436.
40. Baker K, Sengupta D, Salazar-Jimenez G & Cornish VW (2003) An optimized dexamethasone-methotrexate yeast 3-hybrid system for high-throughput screening of small molecule-protein interactions. *Anal Biochem* **315**, 134-137.

41. SenGupta DJ, Zhang B, Kraemer B, Pochart P, Fields S & Wickens M (1996) A three-hybrid system to detect RNA-protein interactions in vivo. *Proc Natl Acad Sci U S A* **93**, 8496-8501.
42. Becker F, Murthi K, Smith C, Come J, Costa-Roldan N, Kaufmann C, Hanke U, Degenhart C, Baumann S, Wallner W, Huber A, Dedier S, Dill S, Kinsman D, Hediger M, Bockovich N, Meier-Ewert S, Kluge AF & Kley N (2004) A three-hybrid approach to scanning the proteome for targets of small molecule kinase inhibitors. *Chem Biol* **11**, 211-223.
43. Dirnberger D, Unsin G, Schlenker S & Reichel C (2006) A small-molecule-protein interaction system with split-ubiquitin as sensor. *Chembiochem* **7**, 936-942.
44. Azizi B, Chang EI, Schwimmer LJ & Doyle DF (2009) *In Progress*.
45. Abida WM, Carter BT, Althoff EA, Lin H & Cornish VW (2002) Receptor-dependence of the transcription read-out in a small-molecule three-hybrid system. *Chembiochem* **3**, 887-895.
46. Althoff EA & Cornish VW (2002) A bacterial small-molecule three-hybrid system. *Angew Chem Int Ed Engl* **41**, 2327-2330.
47. Athavankar S & Peterson BR (2003) Control of gene expression with small molecules: biotin-mediated acylation of targeted lysine residues in recombinant yeast. *Chem Biol* **10**, 1245-1253.
48. Doyle DF, Braasch DA, Jackson LK, Weiss HE, Boehm MF, Mangelsdorf DJ & Corey DR (2001) Engineering orthogonal ligand-receptor pairs from "near drugs". *J Am Chem Soc* **123**, 11367-11371.
49. Horswill AR, Savinov SN & Benkovic SJ (2004) A systematic method for identifying small-molecule modulators of protein-protein interactions. *Proc Natl Acad Sci U S A* **101**, 15591-15596.
50. Muddana SS & Peterson BR (2004) Facile synthesis of cids: biotinylated estrone oximes efficiently heterodimerize estrogen receptor and streptavidin proteins in yeast three hybrid systems. *Org Lett* **6**, 1409-1412.
51. Peet DJ, Doyle DF, Corey DR & Mangelsdorf DJ (1998) Engineering novel specificities for ligand-activated transcription in the nuclear hormone receptor RXR. *Chem Biol* **5**, 13-21.
52. Pollock R & Clackson T (2002) Dimerizer-regulated gene expression. *Curr Opin Biotechnol* **13**, 459-467.
53. Baker K, Blecinski C, Lin H, Salazar-Jimenez G, Sengupta D, Krane S & Cornish VW (2002) Chemical complementation: a reaction-independent genetic assay for enzyme catalysis. *Proc Natl Acad Sci U S A* **99**, 16537-16542.

54. Tao H, Peralta-Yahya P, Decatur J & Cornish VW (2008) Characterization of a new glycosynthase cloned by using chemical complementation. *Chembiochem* **9**, 681-684.
55. Boeke JD, LaCroute F & Fink GR (1984) A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Mol Gen Genet* **197**, 345-346.
56. Allen JB, Walberg MW, Edwards MC & Elledge SJ (1995) Finding prospective partners in the library: the two-hybrid system and phage display find a match. *Trends Biochem Sci* **20**, 511-516.
57. Phizicky EM & Fields S (1995) Protein-protein interactions: methods for detection and analysis. *Microbiol Rev* **59**, 94-123.
58. Rebar EJ, Greisman HA & Pabo CO (1996) Phage display methods for selecting zinc finger proteins with novel DNA-binding specificities. *Methods Enzymol* **267**, 129-149.
59. Smith GP & Petrenko VA (1997) Phage Display. *Chem Rev* **97**, 391-410.
60. Vidal M & Legrain P (1999) Yeast forward and reverse 'n'-hybrid systems. *Nucleic Acids Res* **27**, 919-929.
61. Dmitrova M, Younes-Cauet G, Oertel-Buchheit P, Porte D, Schnarr M & Granger-Schnarr M (1998) A new LexA-based genetic system for monitoring and analyzing protein heterodimerization in Escherichia coli. *Mol Gen Genet* **257**, 205-212.
62. Hu JC, Kornacker MG & Hochschild A (2000) Escherichia coli one- and two-hybrid systems for the analysis and identification of protein-protein interactions. *Methods* **20**, 80-94.
63. Hu JC, O'Shea EK, Kim PS & Sauer RT (1990) Sequence requirements for coiled-coils: analysis with lambda repressor-GCN4 leucine zipper fusions. *Science* **250**, 1400-1403.
64. Joung JK, Chung EH, King G, Yu C, Hirsh AS & Hochschild A (1995) Genetic strategy for analyzing specificity of dimer formation: Escherichia coli cyclic AMP receptor protein mutant altered in its dimerization specificity. *Genes Dev* **9**, 2986-2996.
65. Burz DS, Beckett D, Benson N & Ackers GK (1994) Self-assembly of bacteriophage lambda cI repressor: effects of single-site mutations on the monomer-dimer equilibrium. *Biochemistry* **33**, 8399-8405.
66. Marchetti A, Abril-Marti M, Illi B, Cesareni G & Nasi S (1995) Analysis of the Myc and Max interaction specificity with lambda repressor-HLH domain fusions. *J Mol Biol* **248**, 541-550.

67. Zeng X, Herndon AM & Hu JC (1997) Buried asparagines determine the dimerization specificities of leucine zipper mutants. *Proc Natl Acad Sci U S A* **94**, 3673-3678.
68. Jappelli R & Brenner S (1996) Interaction between cAMP-dependent protein kinase catalytic subunit and peptide inhibitors analyzed with lambda repressor fusions. *J Mol Biol* **259**, 575-578.
69. Dove SL, Joung JK & Hochschild A (1997) Activation of prokaryotic transcription through arbitrary protein-protein contacts. *Nature* **386**, 627-630.
70. Gillies AR, Skretas G & Wood DW (2008) Engineered systems for detection and discovery of nuclear hormone-like compounds. *Biotechnol Prog* **24**, 8-16.
71. Skretas G & Wood DW (2005) Rapid detection of subtype-selective nuclear hormone receptor binding with bacterial genetic selection. *Appl Environ Microbiol* **71**, 8995-8997.

CHAPTER 3

BACTERIAL CHEMICAL COMPLEMENTATION:

ELIMINATING BACKGROUND

3.1 Bacterial Chemical Complementation: Bacteria vs. Yeast

Our lab previously developed chemical complementation (CC), a generalizable selection system in *S. cerevisiae* used to investigate protein-small molecule interactions [1, 2]. CC combines the function of nuclear receptors, ligand activated transcription factors that control gene expression, with the power of genetic selection. This system uses genetic selection to link the survival of a yeast cell to the presence of a specific small molecule [1, 2].

Briefly, CC comprises a human nuclear receptor protein, its co-activator protein, and a small molecule ligand. The system is assembled based on the Y2H (Y2Hs reviewed in Chapter 2.2). Upon binding of its specific small molecule ligand, the nuclear receptor-GAL4 DBD and coactivator-GAL4 AD hybrid proteins associate and activate transcription of a selectable marker [1, 2] (Figure 2.3). The chemical complementation system developed in yeast is highly sensitive (producing growth at 10 nM ligand concentration), has zero background (no growth without ligand) and high dynamic range (growth density and time equivalent to Gal4 induced growth) [2, 3]. This system offers a general method of engineering receptors that activate transcription in response to arbitrary small molecules [4].

Although chemical complementation in yeast has proven to be a very efficient method for screening libraries for functional ligand-receptor pairs, extending chemical complementation to bacteria has several advantages. In principle, a B2H system can

facilitate the rapid analysis of larger libraries than could be evaluated in yeast. Bacteria, specifically *E. coli*, grow faster than yeast, 24 hours vs. 48-72 hours, and have higher transformation efficiency than yeast, 10^{8-9} vs. 10^5 [5, 6]. In addition to the speed and transformation efficiencies of bacteria, some eukaryotic proteins or small molecules may be toxic in yeast. This toxicity is usually a result of these specific proteins or small molecules introduced into the Y2H having the ability to interfere with endogenous homologs in yeast. An excellent example that would also directly affect CC systems is small molecules that have fungicidal activity. These molecules would make yeast an unusable host and the use of a system developed in bacteria essential.

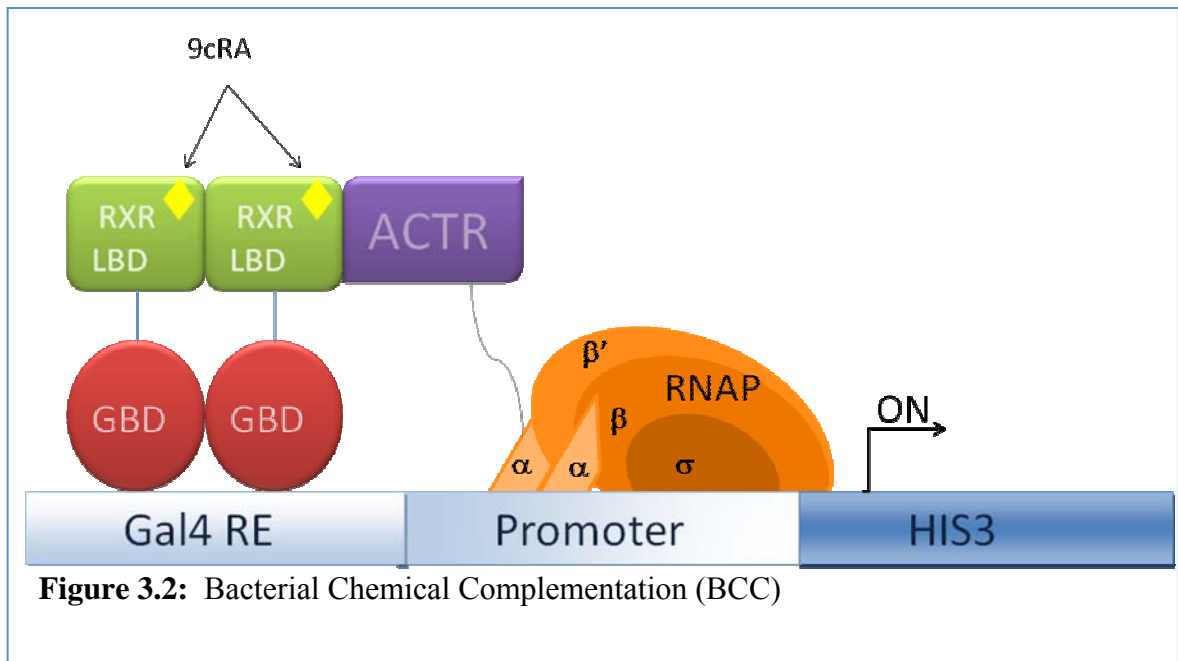
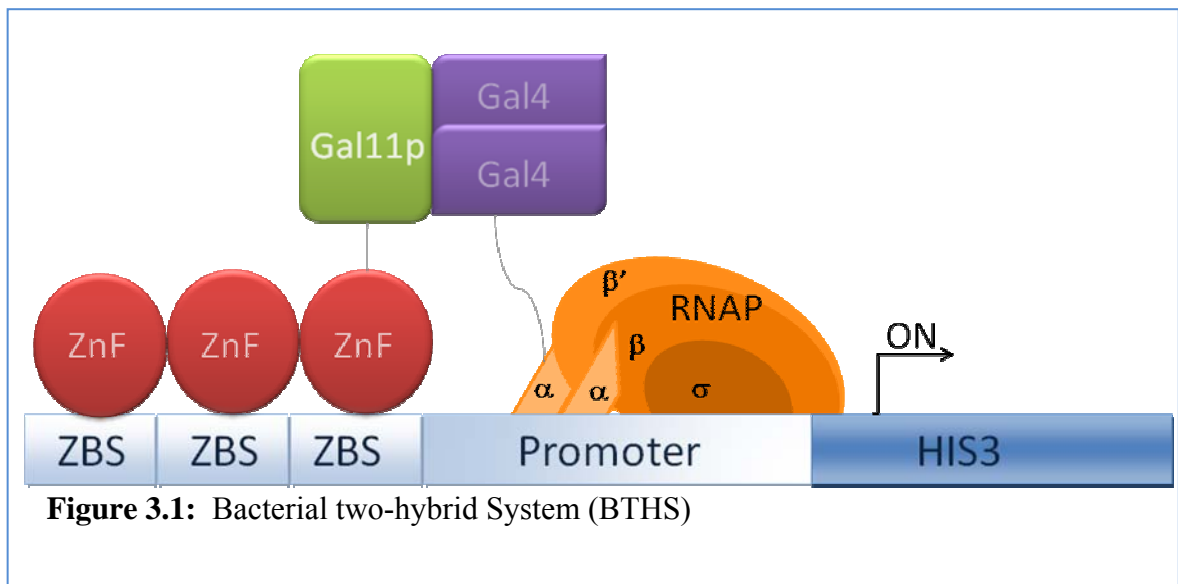
Yeast are simple eukaryotes that contain similar proteins and small molecules to higher eukaryotes, which is one of the appeals to having a yeast based two hybrid system. However, this characteristic may also cause unwanted side effects as well. Introducing higher eukaryote proteins into yeast may result in unintended interactions between the introduced proteins and the endogenous yeast proteins. Bacterial based systems may possible overcome these obstacles. Endogenous interference by small molecules could specifically be an issue for CC in yeast. Some nuclear receptors tested in yeast, particularly the pregnane X receptor (PXR), have been found to be constitutively active. This constitutive activity is hypothesized to be due to endogenous ligand(s) binding the receptor and turning on expression of the selection marker, eliminating our ability to control expression with the desired ligand. Bacteria could eliminate the constitutive activity observed in CC, because they have smaller genomes, less complex proteins, and greater evolutionary distance from higher eukaryotes. Prokaryotes are much simpler in

their protein processing, thus decreasing the endogenous interference seen with simple eukaryotes like yeast.

Most importantly, one of the main advantages to having a bacterial based CC system is that *E. coli* may be a better host for drug screening due to more permeable cell membranes as compared to *S. cerevisiae*, allowing small molecules to cross the membrane and enter the cell easier [7]. Permeability is essential since the CC assay requires the entry of various small molecules into the cell. Based on the B2H developed by Joung and colleagues, a BCC model was devised. This system has the capability to open dimensions for applications such as protein and enzyme engineering, deciphering biosynthetic pathways and drug discovery (explained further in Chapter 3.2).

3.2 Developing the Bacterial Chemical Complementation

Joung and colleagues engineered a bacterial two-hybrid genetic selection system (BTHS) controlling a *HIS3* selective gene (Figure 3.1) Similar to the yeast two-hybrid systems, this system works via the interaction of two fusion proteins that leads to transcription of an essential gene. The *HIS3* gene is a selection marker which encodes for the enzyme imidazole-glycerolphosphate dehydratase. This enzyme is crucial to histidine biosynthesis and allows cells lacking a functional *HIS3* gene to grow in media lacking histidine. Joung began by constructing an *E. coli* strain, KJIC, which contains a deletion in the *hisB* gene, a bacterial gene essential for histidine biosynthesis and cell survival. Joung's BTHS was constructed using two fusion proteins; the human Zif268 DBDs fused to Gal11P protein and the Gal4 proteins dimerization domain (residues 50-97) fused to the alpha subunit of RNA polymerase. Gal11P is a mutant of the Gal11 protein, a component of the RNA polymerase II holoenzyme, where P stands for transcriptional



potentiator [8]. The Gal4 dimerization domain not only interacts with Gal11P, but is also a powerful transcriptional activator in *S. cerevisiae* cells bearing Gal11P [9-11]. This system uses the interaction between Gal11P and Gal4 to regulate expression of a *HIS3* gene [5].

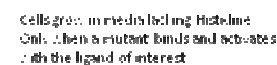
The use of the *HIS3* selection marker as opposed to a screening marker, such as the *lacZ* gene, not only allows one to employ the power of genetic selection, but also allows one to evaluate libraries up to 10^{8-9} vs. 10^{5-6} in size [5]. The novelty of the BTHS is that it can be used to not only evaluate protein-protein interactions, but protein-DNA, and potentially protein-small molecule, interactions as well. Joung and colleagues tested the BTHS by selecting zinc finger variants, from a large randomized library, that bind specific DNA targets in a single selection step [5]. We proposed to adapt this system so that mammalian nuclear receptors could be evaluated for their ability to activate transcription in bacteria in response to a specific ligand. This system would be analogous to CC in yeast.

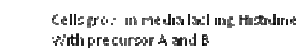
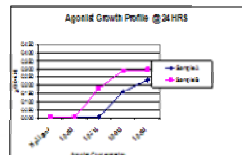
To directly parallel the *S. cerevisiae* CC system, we adapted the Joung Zif268 DBD-Gal11P fusion protein to contain the Gal4 DBD fused to the nuclear receptor RXR. Additionally, a known coactivator of RXR, either the activator of thyroid and retinoic acid *receptor* (ACTR) or the steroid receptor *coactivator-1* (SRC-1), was fused to the alpha subunit of RNA polymerase (RNAP). The reporter construct was also adapted to contain a Gal4 response element upstream of the lac promoter (P_{wk}) and *HIS3* gene. Both fusion proteins and the reporter construct were introduced into the *E. coli* strain KJIC and grown in histidine selective media with and without 9-*cis* retinoic acid (9cRA), the natural ligand for RXR.

Ideally, the Gal4 DBD will bind the Gal4 response elements (REs) and in the presence of the appropriate ligand, 9cRA, the ligand will bind the nuclear receptor LBD. The LBD will undergo a conformation change, and recruit the coactivator-alpha RNA polymerase fusion protein. The alpha subunit of RNA polymerase associates with the RNA polymerase core enzyme (RNAP), recruiting the entire RNAP to the promoter and activating transcription (Figure 3.2). This system would allow us to create a generalizable selection system in bacteria that could potentially be used for drug discovery, protein and enzyme engineering, as well as deciphering biosynthetic pathways.

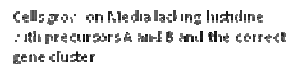
Applications of Bacterial Chemical Complementation

As mentioned above, BCC can be used as an essential tool in a variety of applications. To date, CC has been used as a very effective tool for protein engineering and analyzing NR libraries [4]. BCC, like CC, can not only be used as a very efficient method to analyze potentially large protein libraries, but will also allow us to obtaining insight into NR structure and function. When using BCC for protein engineering, NR libraries are rationally designed for a specific ligand and analyzed using BCC (Figure 3.3A). In addition to protein engineering, BCC can be used as a high throughput assay for drug discovery. As mentioned in Chapter 1, NRs are involved in a number of diseases and small molecules which regulate these receptors make up 10% of pharmaceuticals today. In this application, chemical complementation is used to analyze libraries of small molecules to select for those able to bind and activate a specific nuclear receptor. Because BCC is a selection assay, one can analyze many compounds at one time and only those which are functional will have to be evaluated, as opposed to a

Library of Mutated
Nuclear Receptors



Biosynthetic Pathways



screening system in which all variants, functional and nonfunctional, are evaluated (Figure 3.3B). For this assay a 96 well plate format can be used in which each well encompasses cells containing the BCC system and a different small molecule. To select for potential agonists, only wells containing a small molecule able to bind and activate the nuclear receptor will grow in selective media.

BCC can serve as a powerful tool for enzyme engineering as well. This process involves creating libraries of enzyme variants capable of converting particular substrates into the enzymatic product of interest. The production of the desired enzymatic product can be analyzed using BCC, opening a new dimension of assaying enzymatic libraries. If the enzymatic product is produced, the product will bind the nuclear receptor associated with BCC and activate transcription of an essential gene. This system requires one to have a nuclear receptor which has already been discovered, or engineered, to bind the enzymatic product. This application, like protein engineering, would also give insight into the structure and function limitations of enzyme engineering (Figure 3.3C).

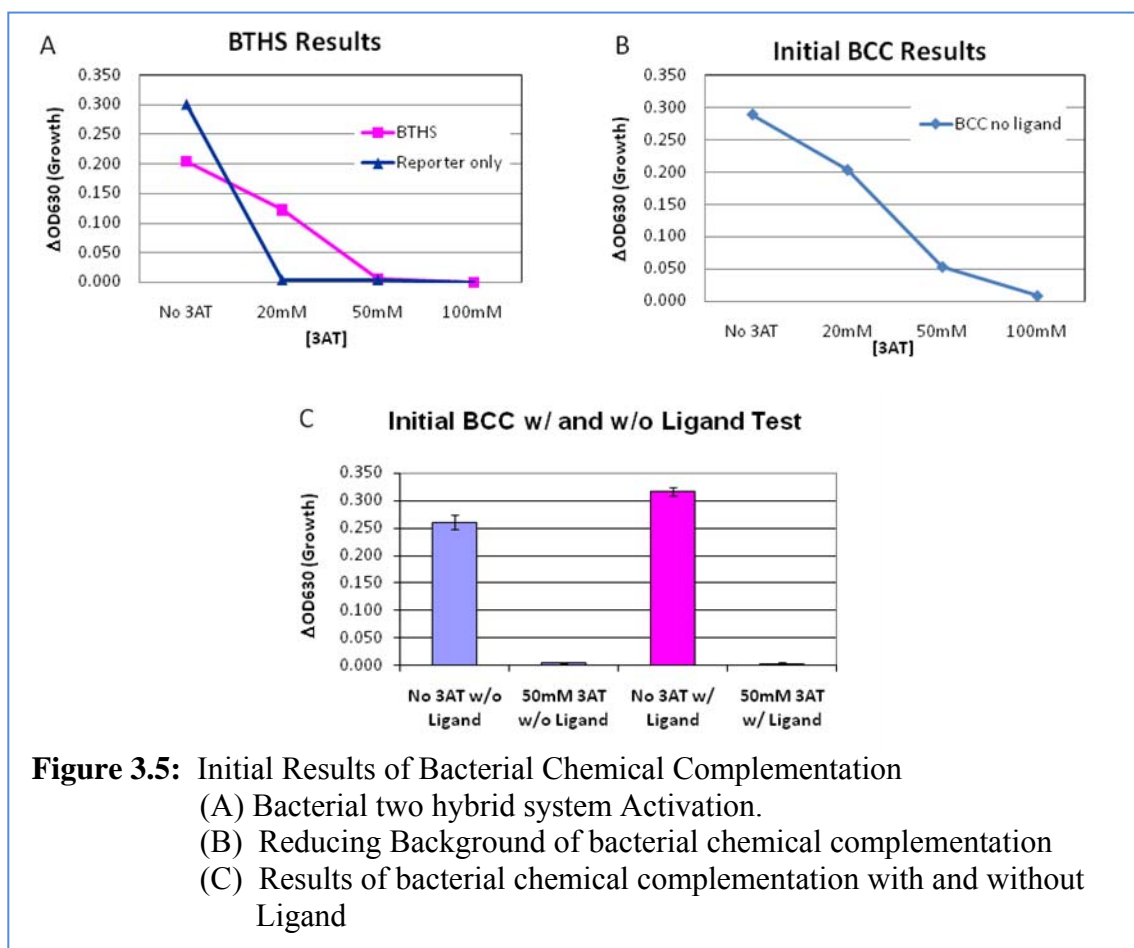
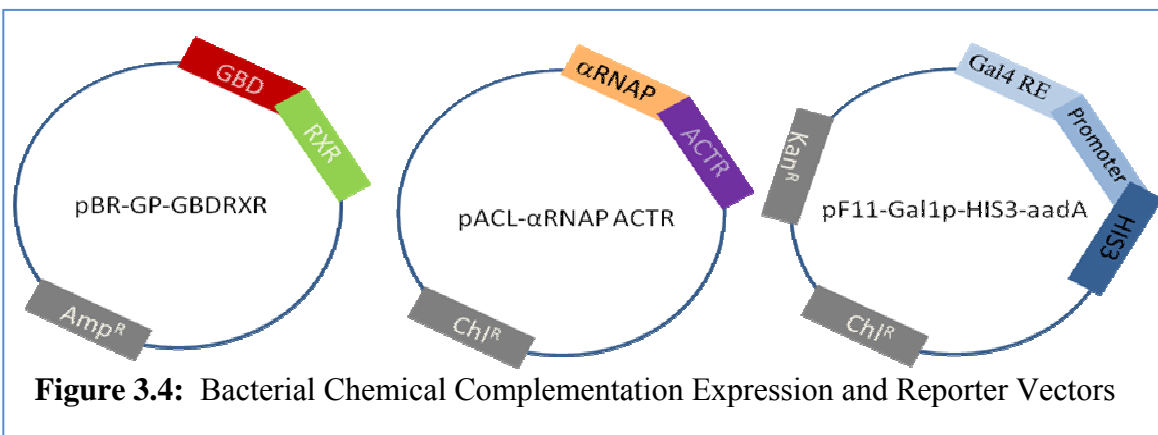
Similarly to enzyme engineering, BCC can be used to decipher biosynthetic pathways as well. Natural products account for one of the major sources for drug discovery by pharmaceutical companies [12]. An alternative approach to chemical synthesis of these natural products is through identification of biosynthetic pathways used to produce these natural compounds. In order to biosynthetically produce these natural products, one must be able to identify the biosynthetic genes involved in producing these compounds from their natural source. Once identified, these genes can be introduced into a heterologous organism and serve as a “factory” for an alternative means of obtaining large amounts of these compounds.

To use BCC to decipher through possible genes involved in the biosynthesis of a particular natural product, the first step must be to engineer a nuclear receptor that will bind and activate transcription in response to the natural product of interest. The next step would be to create cDNA libraries of possible biosynthetic gene clusters that could potentially convert the precursors into the desired product. Cells containing BCC and the correct gene cluster will be able to convert the precursors into the product of interest, allowing cells to grow in media lacking histidine (Figure 3.3D). Using BCC in these applications has the ability to open new dimensions in protein and enzyme engineering.

Initial Results of Bacterial Chemical Complementation

The expression and reporter vectors obtained from Joung, pBR-GP-Z123, pACL- α Gal4, and pF11-Z321-HIS3-aadA [5], were modified to create the expression and reporter vectors needed for bacterial chemical complementation. The expression plasmid/vector, pBR-GP-GBDRXR, contains the GBD-RXR fusion protein and an ampicillin resistance gene. The pACL- α RNAP ACTR expression plasmid was designed to contain the alpha RNAP subunit-ACTR fusion protein and a chloramphenicol resistance marker. The pF11-Gal1p-HIS3-aadA reporter vector contained the Gal4 response element upstream of weak lactose promoter (Pwk) and the *HIS3* gene and contained both chloramphenicol and kanamycin resistant genes (Figure 3.4).

As a control, the BTHS, containing the bacterial strain, KJIC F' Z321 , and the expression vectors supplied by Joung, were successfully reproduced from the Joung manuscript [5] (Figure 3.5A). Once the BTHS controls were reproduced, BCC expression and reporter vectors were introduced (transformed) into the strain, KJIC. The BCC cells were grown in media lacking histidine, without ligand, to determine the



optimal concentration of 3-amino-1,2,4-triazole (3AT) needed to reduce the background caused by “leaky” expression of the *HIS3* gene. The *HIS3* selective gene encodes imidazole-glycerolphosphate dehydratase. In the absence of activation, there is low constitutive expression of the *HIS3* gene from the weak promoter. This “leaky” background growth can be eliminated with the use of 3-amino-1,2,4-triazole (3AT), an inhibitor of imidazole-glycerolphosphate dehydratase [13]. Joung’s BTHS required 20 mM 3AT to reduce the “leaky” expression [5] and yeast chemical complementation system requires 5 mM 3AT [14].

Initially, 50 mM 3AT was needed to eliminate the background growth observed with the BCC system (growth observed without the presence of ligand) (Figure 3.5B). However, once the background growth was eliminated using 50 mM 3AT, the bacterial chemical complementation system based on RXR and 9cRA revealed no ligand-activated growth occurred above background growth (Figure 3.5C). A functional BCC system requires (1) no background (discussed in the remainder of this chapter) and (2) ligand activated growth (discussed in Chapter 4).

3.3 Results of Eliminating Bacterial Chemical Complementation Background

Initial results from BCC revealed that the system required 50 mM 3AT to reduce background growth from the *HIS3* gene in media lacking histidine. The use of 50 mM 3AT to reduce BCC background was hypothesized to possibly be too stringent, especially in comparison to the BTHS which only requires 20 mM 3-AT [5]. This high background growth observed by the BCC system may indeed represent the maximal growth which can be observed by this system, with or without ligand, and any activation, or growth, observed as a result of the ligands activity cannot be seen. In other words, the

background growth may represent the maximum threshold of activation capable. As a result of this, we hypothesized that before we could see ligand activated growth, we must first find the cause of and reduce the high background observed with BCC.

3.3.1 Effects of Expression Vectors on Background Activation

We began investigating the cause of the high background seen in BCC by looking at each component of the BCC system to see if it was somehow causing expression of the *HIS3* gene in the absence of ligand. To test the possibility that one of the vectors themselves may be producing a hybrid protein that is interfering with endogenous proteins, activating transcription, and causing the high background, both the BCC system and Joung's BTHS were tested with each of the expression vectors, one at a time, and the reporter vector (Figure 3.6A). Every system, whether it was full system, only the reporter, or the reporter with only one of the hybrid proteins, exhibited a high background, both in the BCC system and the BTHS (Figure 3.6B). The reporter vector alone produced a high background, regardless of whether the expression vectors were present or not suggesting that the background growth was most likely not associated with the expression vectors and/or their resulting hybrid proteins, but rather the background may be a result of a problem with the strain, or reporter construct itself.

3.3.2 Effects of Promoters on Background

Since our results concluded that the background growth was not associated with the expression plasmids, but did indicate that the reporter construct plays a role in the high background (Figure 3.6), we hypothesized that the lactose promoter used upstream the *HIS3* gene in our reporter may be the cause of the background. In BCC, the controllable activation of the selective gene is dependent upon having a weak promoter.

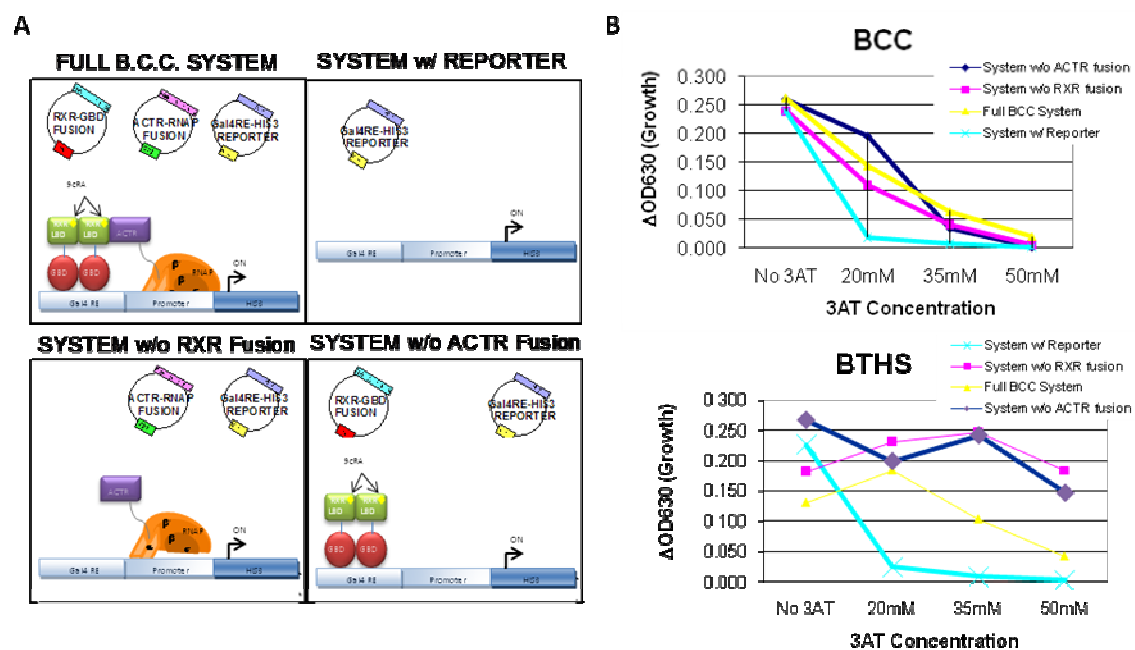


Figure 3.6: Effects of Expression and Reporter Vectors on Background Growth.
Vector systems evaluated Results of vector system on background growth

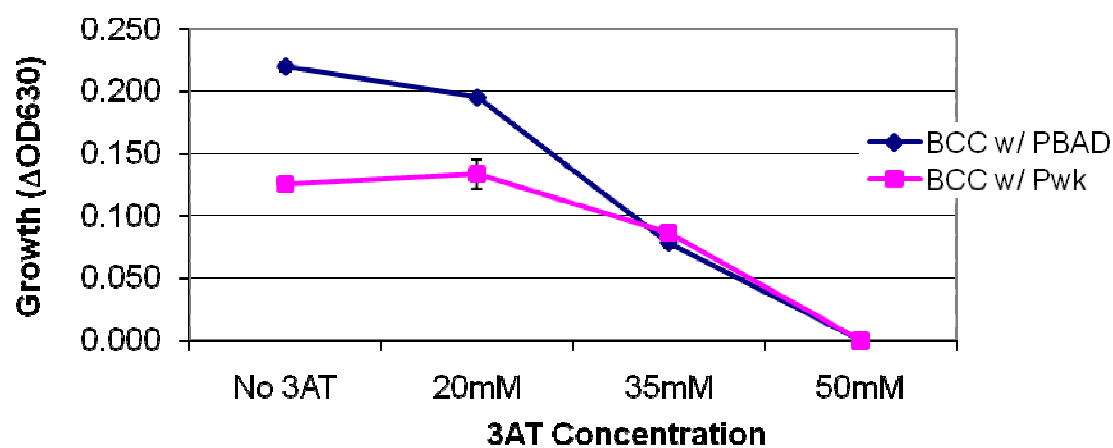


Figure 3.7: Effect of BCC Promoters on Background Growth

When using a weak promoter, the affinity of RNAP for the promoter is very weak, resulting in low levels of activation. However, if the system is able recruit the RNAP core enzyme into the vicinity of the promoter, it will enhance activation. If the promoter is a strong promoter, the RNAP will have a greater affinity for the promoter and not need to be recruited by our system in order to activate, causing enhanced expression of the selective gene, which in turn would cause the selective gene to be expressed even in the absence of a functional two-hybrid system resulting in a high background. We investigated whether the strength of the lac promoter was causing our high background by replacing the lactose promoter with another known bacterial weak promoter, the arabinose promoter.

The lactose promoter is a component of the lac operon, which is responsible for lactose metabolism and transport in *E. coli*. The regulation of this operon, by the availability of glucose and lactose, was the first genetic regulatory mechanism elucidated, and is also a popular tool used by two-hybrid assays. The lactose operon is one of the most intensely studied genetic regulatory systems and is extremely useful in genetic engineering. In order for RNAP to effectively bind the lac promoter and induce transcription of the downstream gene, a positive activator called the catabolite activator protein (CAP) must binds to a specific sequence immediately next to the promoter which helps RNAP to bind and activate [15]. The lac promoter, associated with the lac operon, is considered a weak promoter in that its -35 and -10 consensus sequences differ from the ideal -35 and -10 sequences and as a result RNA polymerase (RNAP) does not bind well to it. Although many sources describe the lac promoter as a weak promoter, there are also sources which also describe the [lac](#) promoter as a medium strength promoter, and as

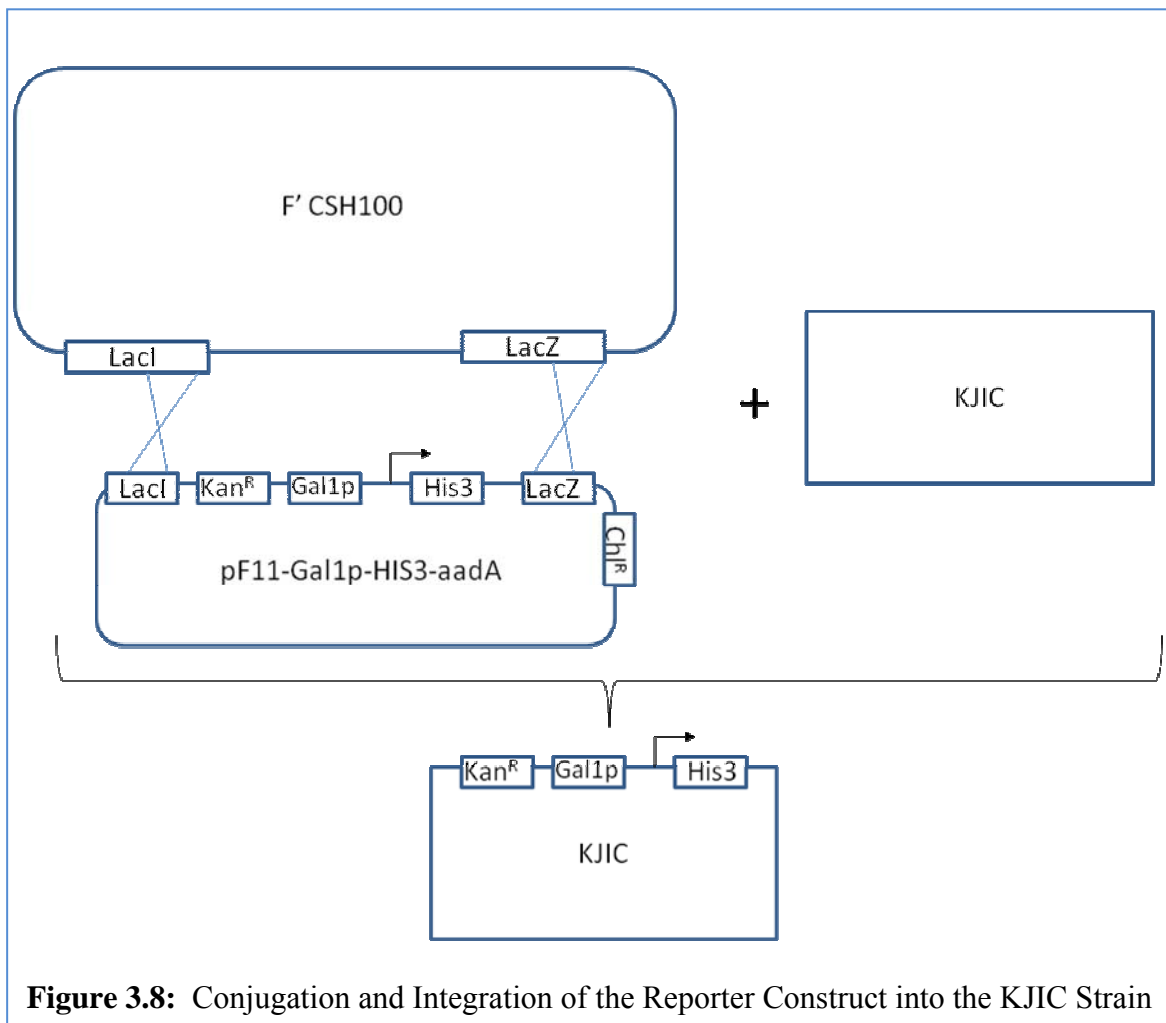
a result the amount of mRNA made and, hence, the amount of the protein products will be moderate.

The arabinose operon is responsible for producing enzymes necessary for arabinose metabolism, and like the lac operon, is regulated by a dual positive and negative system via the use of arabinose. The arabinose promoter is often used as an alternative to the lac promoter and has become very popular for controlling gene expression in *E. coli* [16]. This weak but tightly controlled promoter–regulator system has enabled growth of both soluble and insoluble heterologous proteins in high-cell-density *E. coli* cultures [17]. For this reason, we decided to replace the lac promoter (P_{wk}) with the arabinose promoter (P_{BAD}).

When comparing BCC using the arabinose promoter vs. the lactose promoter we found that both systems worked fairly similarly. Both required 50 mM 3AT to reduce the background observed by the *HIS3* gene in the absence of ligand (Figure 3.7). If the arabinose promoter was indeed a weaker promoter than the lac promoter it should require less 3AT to reduce the background. However, our results indicate that the promoter is not the cause of the high background observed with BCC.

3.3.3 Effects of Integrating the Reporter on Background Growth

Our system, unlike the system reported in the Joung manuscript, used the reporter construct in vector form as opposed to integrated into the genome of the strain. Using the unintegrated reporter in the BCC system, would result in a much higher vector copy number, as opposed to a single copy obtained when the construct is integrated into the strain. Hence, the additional copies of the *HIS3* gene would be present in the cell, which are all leaky, resulting in a much higher background in the absence of



ligand than would be observed if only a single copy of the reporter construct were present.

To investigate whether integration of the reporter construct would reduce the high background we were observing, the reporter construct, containing the *HIS3* gene under the control of a lactose promoter and the Gal4 promoter (Gal1p, containing two Gal4 response elements), was conjugated into the F' episome strain CSH100 and integrated into the KJIC strain using the method outlined by Whipple and colleagues [18] to obtain an *E. coli* KJIC strain referred to as KJIC F' Gal1p 2RE. Briefly, the conjugation method involves transferring the reporter construct from a customized plasmid onto an F' episome in the CSH100 strain via homologous recombination. The construct is then moved from the F' episome to the KJIC strain through direct cell-to-cell contact, referred to as conjugation (Figure 3.8). We then tested both the integrated BCC system, as well as the, integrated BTHS (this same system was previously tested in Figure 3.6A) to see if we were able to reduce the high background observed in the integrated system.

We found that in the integrated systems, both the BTHS as well as the BCC required 20 mM 3AT to reduce the background caused by leaky expression of the *HIS3* gene (Figure 3.9). After successfully reducing the high background, we went on to test whether or not we could obtain ligand activated growth with the conjugated BCC system.

3.4 Materials and Methods

Bacterial Strains

E. coli strains, CS H100 (F' lac proA⁺B⁺ [lacI^q lacPL8]/ara- Δ[gpt-lac]5), KJIC (F-ΔhisB463 Δ[gpt-proAB-arg-lac] XIII zaj::Tn10), and KJIC F' Z321 were a kind gift from Dr. Keith Joung (Massachusetts General Hospital/Harvard Medical School). The

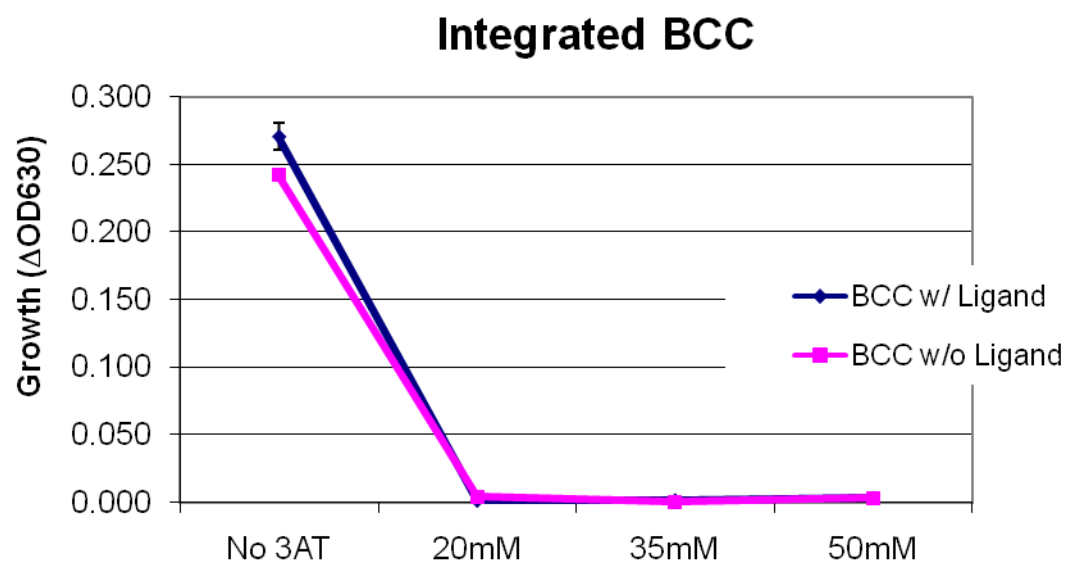
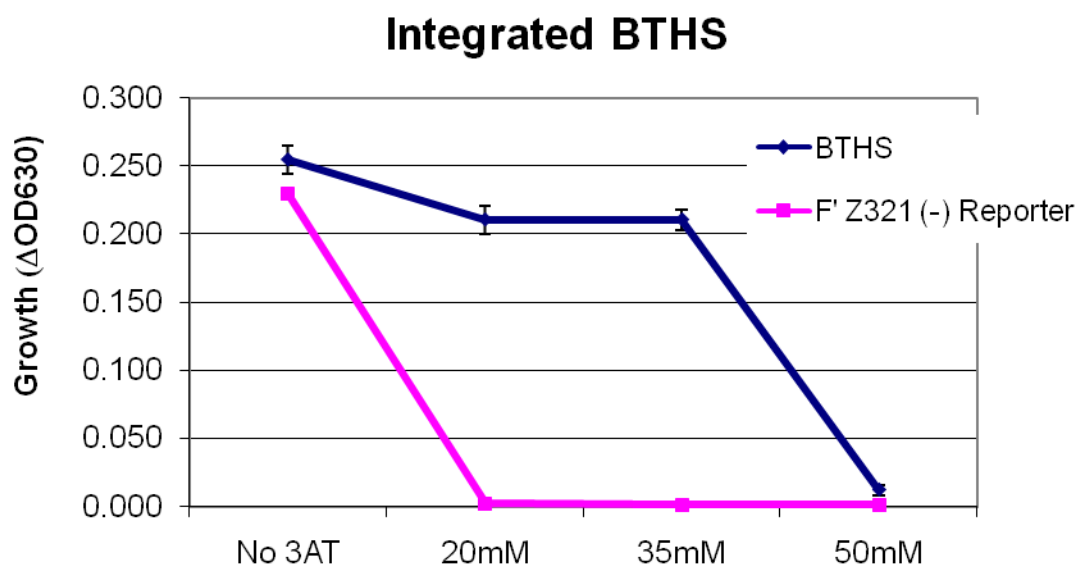


Figure 3.9: Effects of Reporter Construct Integration on Background Growth

KJIC F' Z321 strain contains the Zif268 binding site upstream of the transcription start site of a weak *E. coli* lac-promoter and controlling the *HIS3* gene [5, 18]. The KJIC F' Gal1p 2RE was constructed by conjugating the Gal1p-2RE-Pwk-His3-aadA construct from pF11-Gal1p-HIS3-aadA into the KJIC strain, via the CSH100 strain, using a previously described protocol [18].

Ligands

9 cis-retinoic acid (MW=300.44 g/mol) was purchased from Biomol (Plymouth Meeting, PA). 10 mM stocks of the ligand were dissolved in 80% ethanol:20% DMSO (4:1 v/v) and stored at 4°C protected from light.

Expression plasmids and Reporter Constructs

pBR-GP-Z123, pACL- α Gal4, and pF11-Z321-HIS3-aadA were kind gifts from Dr. Keith Joung (Massachusetts General Hospital/Harvard Medical School) [5]. To make pBR-GP-GBDRXR, the restriction site, HindIII, was inserted into the pBR-GP-Z123 vector at the 5' end of the zinc finger proteins, Z123, using site directed mutagenesis. Primers containing a HindIII site at the 5' end and a AatII site at the 3' end were designed to amplify the GBDRXR fusion gene (containing residues 1-151 of Gal4 and residues 44-462 of RXR) from the pGBDRXR vector (previously constructed in our lab) and ligate it into the pBR-GP-Z123-HindIII vector using the HindIII and AatII restriction enzymes.

To make pACL- α RNAP ACTR, primers containing a NotI site at the 5' end and a AvrII site at the 3' end were designed to amplify the ACTR gene (from start to stop) from the pGAD10BA ACTR vector (previously constructed in our lab) and ligate it into the pACL- α Gal4 vector using the NotI and AvrII restriction enzymes.

To make pF11-Gal1p-HIS3-aadA, primers containing a EagI site at both the 5' and 3' ends were designed to amplify the Gal1p response elements (containing 2 response elements) from the pGH1 vector, a kind gift from Dr. Philip James (University of Wisconsin, Madison), and ligated into the pF11-Z321-HIS3-aadA vector using the EagI restriction enzyme. To make pF11-Gal1p-AraBAD-HIS3-aadA, the weak lac promoter (TTTACCA-18bps-TATGTT) was replaced with the arabinose promoter (CTGACG-18bps-TACTGT) via site directed mutagenesis.

After digestion and standard ligation, each vector was transformed into Zcompetent (Zymo Research, Orange, CA) XL1-Blue cells and selected for growth onto LB plates with the appropriate antibiotics (50 mg/ml Ampicillin, 25 µg/ml of chloramphenicol, or 30 µg/ml kanamycin, respectively). The resulting vectors were purified using QIAprep Spin Miniprep Kit (Qiagen Inc., Valencia, CA), diagnosed with restriction enzymes and sequenced for confirmation.

Media

Luria-Bertani Media (LB) media was made by dissolving 25 g of Luria broth powder in 1 L of distilled water. The pH was adjusted to 7.0 with either HCl or NaOH as appropriate and sterilize by autoclaving. M9-His selective media was made according to the BacterioMatch® II Two-Hybrid System Vector Kit (Stratagene, La Jolla, CA)) with 9cRA (10 µM) and antibiotics (50 mg/ml Ampicillin (Amp), 25 µg/ml of chloramphenicol (Chl), and 30 µg/ml kanamycin (Kan)). 3-amino-1, 2,4-triazole (3-AT) was purchased from Sigma (St. Louis, MO), dissolved in water to make 100 mM stock solutions, filter sterilized and stored at room temperature.

Bacterial Selection Protocols

The BCC systems (consisting of pBR-GP-GBDRXR, pACL- α RNAP ACTR, and pF11-Gal1p-HIS3-aadA vectors) and the BTHS (consisting of pBR-GP-Z123, pACL- α Gal4, and pF11-Z321-HIS3-aadA) were transformed into Zcompetent (Zymo Research, Orange, CA) KJIC cells and selected for growth onto LB plates with the appropriate antibiotics, 50 mg/ml Ampicillin (Amp), 25 μ g/ml of chloramphenicol (Chl), and/or 30 μ g/ml kanamycin (Kan), respectively. Transformants were selected and grown in LB Amp, Chl, Kan media overnight. Transformed cells were washed and re-inoculated into M9-His selective with or without 9cRA with the appropriate antibiotics and plated in 96 well plates. Cells were grown at 37 °C at 300 rpm for 24 hours to select for ligand activated growth.

Expression vector systems (containing the pF11-Gal1p-HIS3-aadA vector and either the pBR-GP-GBDRXR or the pACL- α RNAP ACTR for BCC systems or pF11-Z321-HIS3-aadA and either pBR-GP-Z123 or pACL- α Gal4 for BTHS) were transformed like above but selected for with either Kan/Amp media or Kan/Chl media, respectively. The conjugated BTHS and BCC systems (consisting of pBR-GP-Z123 and pACL- α Gal4 and pBR-GP-GBDRXR and pACL- α RNAP ACTR, respectively) transformed into Zcompetent (Zymo Research, Orange, CA) KJIC F' Z321 or KJIC F' Gal1p 2RE cells, respectively. Transformants were selected for growth in minimal media lacking histidine (M9-His selective media made according to the BacterioMatch® II Two-Hybrid System Vector Kit (Stratagene, La Jolla, CA)) with the appropriate antibiotics either with or without 9cRA as described above.

3.5 Literature Cited

1. Azizi B, Chang EI & Doyle DF (2003) Chemical complementation: small-molecule-based genetic selection in yeast. *Biochem Biophys Res Commun* **306**, 774-780.

2. Azizi B, Chang EI, Schwimmer LJ & Doyle DF (2009) *In Progress*.
3. Azizi B (2005) Chemical complementation a genetic selection system in yeast for drug discovery, protein engineering, and for deciphering and assembling biosynthetic pathways. In *unpublished thesis (PhD)*. Georgia Institute of Technology, Atlanta.
4. Schwimmer LJ, Rohatgi P, Azizi B, Seley KL & Doyle DF (2004) Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules. *Proc Natl Acad Sci U S A* **101**, 14707-14712.
5. Joung JK, Ramm EI & Pabo CO (2000) A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. *Proc Natl Acad Sci U S A* **97**, 7382-7387.
6. Serebriiskii IG, Fang R, Latypova E, Hopkins R, Vinson C, Joung JK & Golemis EA (2005) A combined yeast/bacteria two-hybrid system: development and evaluation. *Mol Cell Proteomics* **4**, 819-826.
7. Vaara M (1992) Agents that increase the permeability of the outer membrane. *Microbiol Rev* **56**, 395-411.
8. Barberis A, Pearlberg J, Simkovich N, Farrell S, Reinagel P, Bamdad C, Sigal G & Ptashne M (1995) Contact with a component of the polymerase II holoenzyme suffices for gene activation. *Cell* **81**, 359-368.
9. Hidalgo P, Ansari AZ, Schmidt P, Hare B, Simkovich N, Farrell S, Shin EJ, Ptashne M & Wagner G (2001) Recruitment of the transcriptional machinery through GAL11P: structure and interactions of the GAL4 dimerization domain. *Genes Dev* **15**, 1007-1020.
10. Dove SL & Hochschild A (1998) Conversion of the omega subunit of Escherichia coli RNA polymerase into a transcriptional activator or an activation target. *Genes Dev* **12**, 745-754.
11. Farrell S, Simkovich N, Wu Y, Barberis A & Ptashne M (1996) Gene activation by recruitment of the RNA polymerase II holoenzyme. *Genes Dev* **10**, 2359-2367.
12. Newman DJ & Cragg GM (2007) Natural products as sources of new drugs over the last 25 years. *J Nat Prod* **70**, 461-477.
13. Struhl K & Davis RW (1977) Production of a functional eukaryotic enzyme in Escherichia coli: cloning and expression of the yeast structural gene for imidazole-glycerolphosphate dehydratase (his3). *Proc Natl Acad Sci U S A* **74**, 5255-5259.
14. Azizi B, Chang EI, Schwimmer LJ & Doyle DF (2005) *Submitted*.

15. Czarniecki D, Noel RJ, Jr. & Reznikoff WS (1997) The -45 region of the Escherichia coli lac promoter: CAP-dependent and CAP-independent transcription. *J Bacteriol* **179**, 423-429.
16. Guzman LM, Belin D, Carson MJ & Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177**, 4121-4130.
17. Lim HK, Jung KH, Park DH & Chung SI (2000) Production characteristics of interferon-alpha using an L-arabinose promoter system in a high-cell-density culture. *Appl Microbiol Biotechnol* **53**, 201-208.
18. Whipple FW (1998) Genetic analysis of prokaryotic and eukaryotic DNA-binding proteins in Escherichia coli. *Nucleic Acids Res* **26**, 3700-3706.

CHAPTER 4

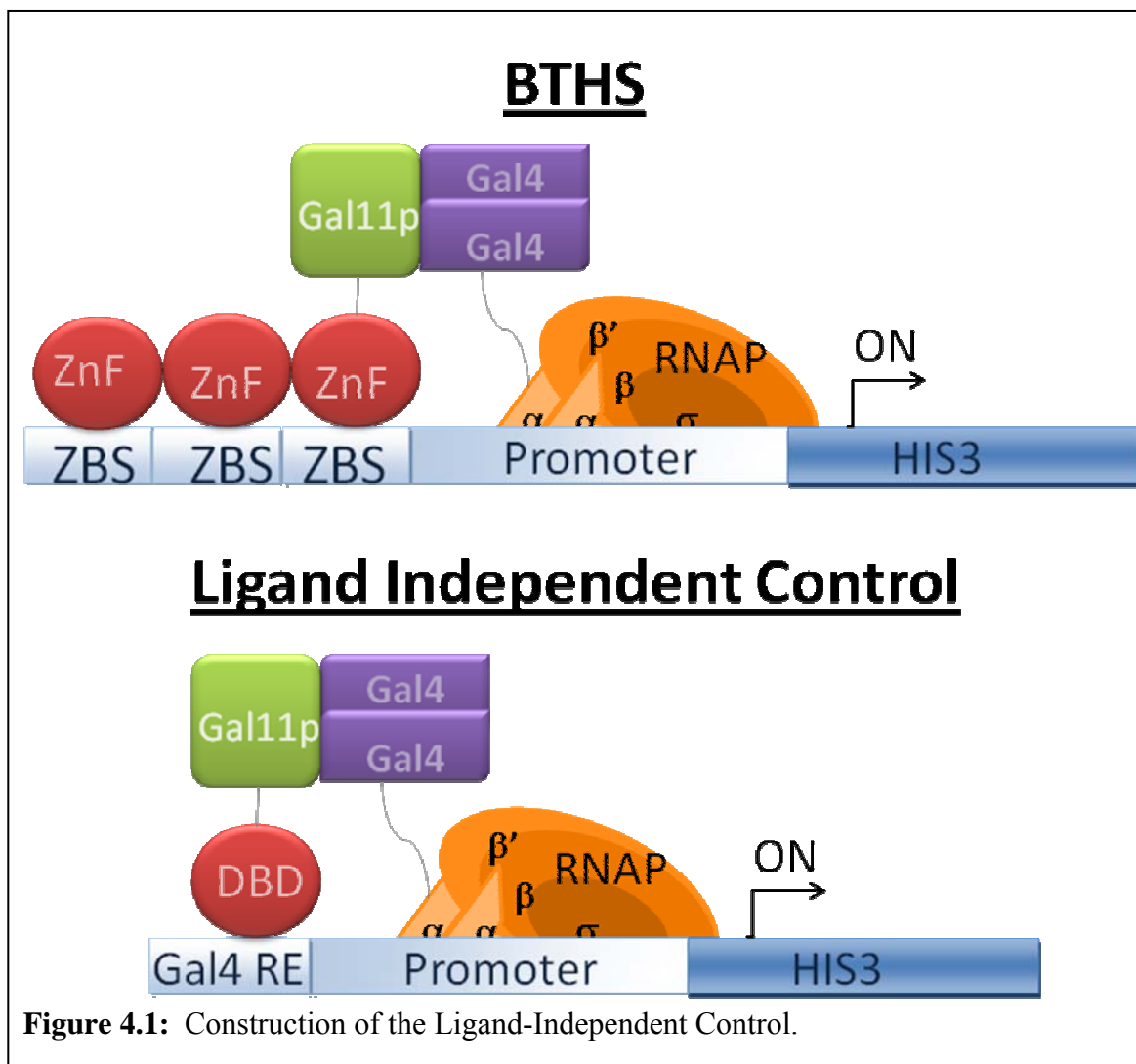
BACTERIAL CHEMICAL COMPLEMENTATION: OBTAINING LIGAND DEPENDENT ACTIVATION

4.1 Optimizing Bacterial Chemical Complementation for Ligand Activation

After conjugating and integrating the Gal1p-*HIS3* reporter construct into the KJIC strain, we were successfully able to reduce the high background initially observed with BCC when no ligand was present. Despite the reduction in background growth, we were still unable to observe ligand activated growth above background (Figure 3.9). Given that both the BTHS and the yeast chemical complementation system were functional, we conclude that the selection assay itself, as well as the interactions between the nuclear receptor and coactivator, should be capable of producing ligand activated growth. To troubleshoot the lack of ligand activated growth, we compared BCC to both the BTHS and the yeast CC to determine the cause of the lack of activation. The remainder of this chapter will discuss the various nuclear receptor-ligand pairs and coactivators tested, the optimization of the Gal4 DBDs interaction with its response elements, and attempts to optimize the coactivator and nuclear receptor heterologous expression and folding to attain ligand dependent activation of BCC.

4.1.1 Enhancing the Interaction Between GBD and Gal1p

To determine whether the Gal4 DBD was successfully capable of binding to the Gal1p response element and enhancing transcription of an essential gene, a control system was constructed, similar to the BTHS control system, requiring an optimal GBD



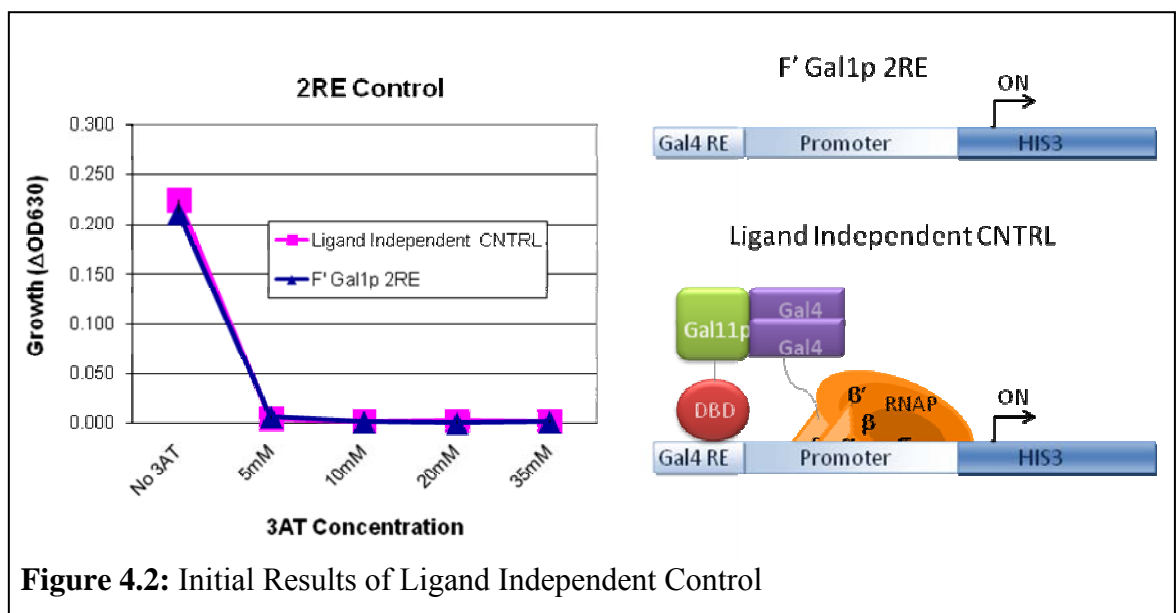
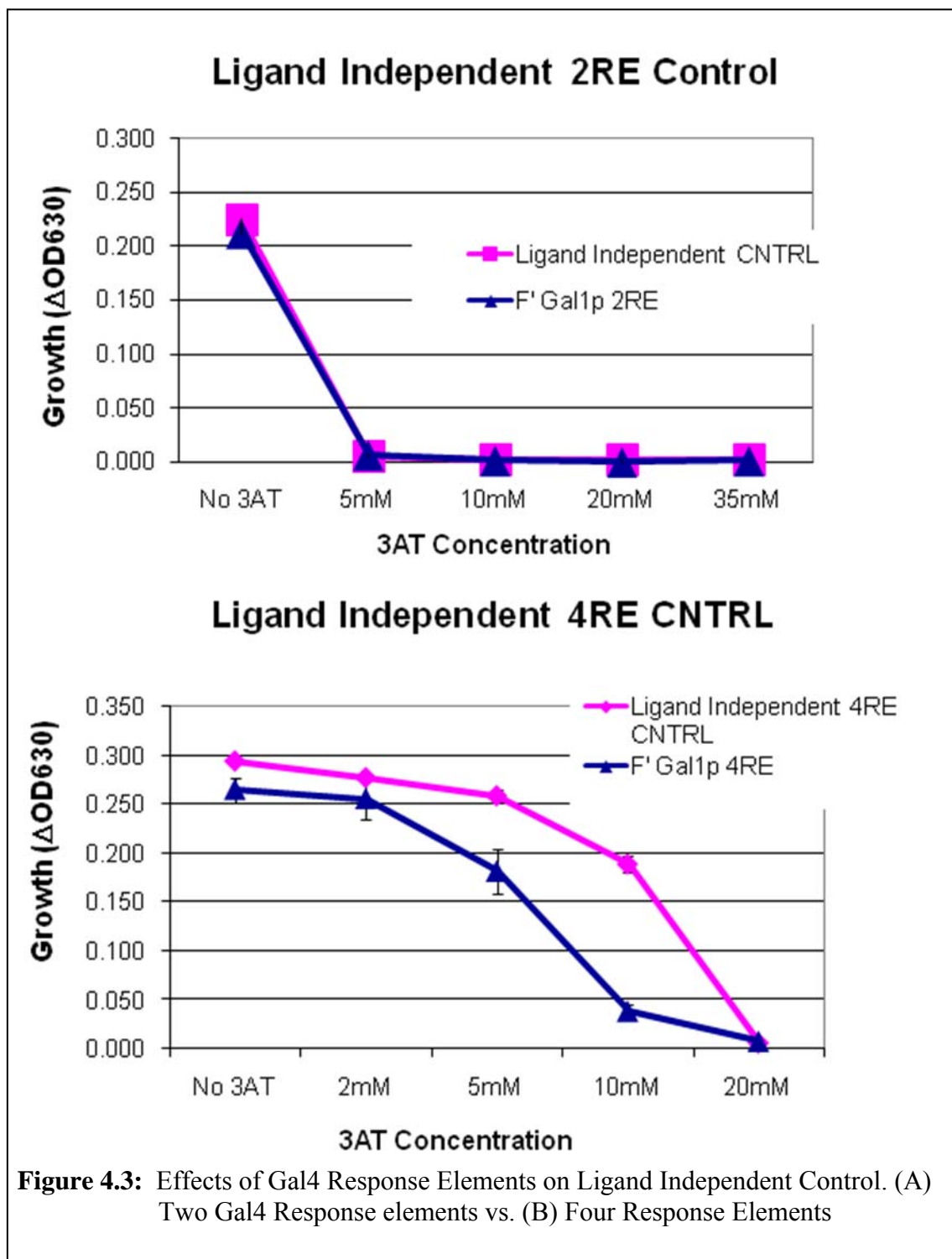


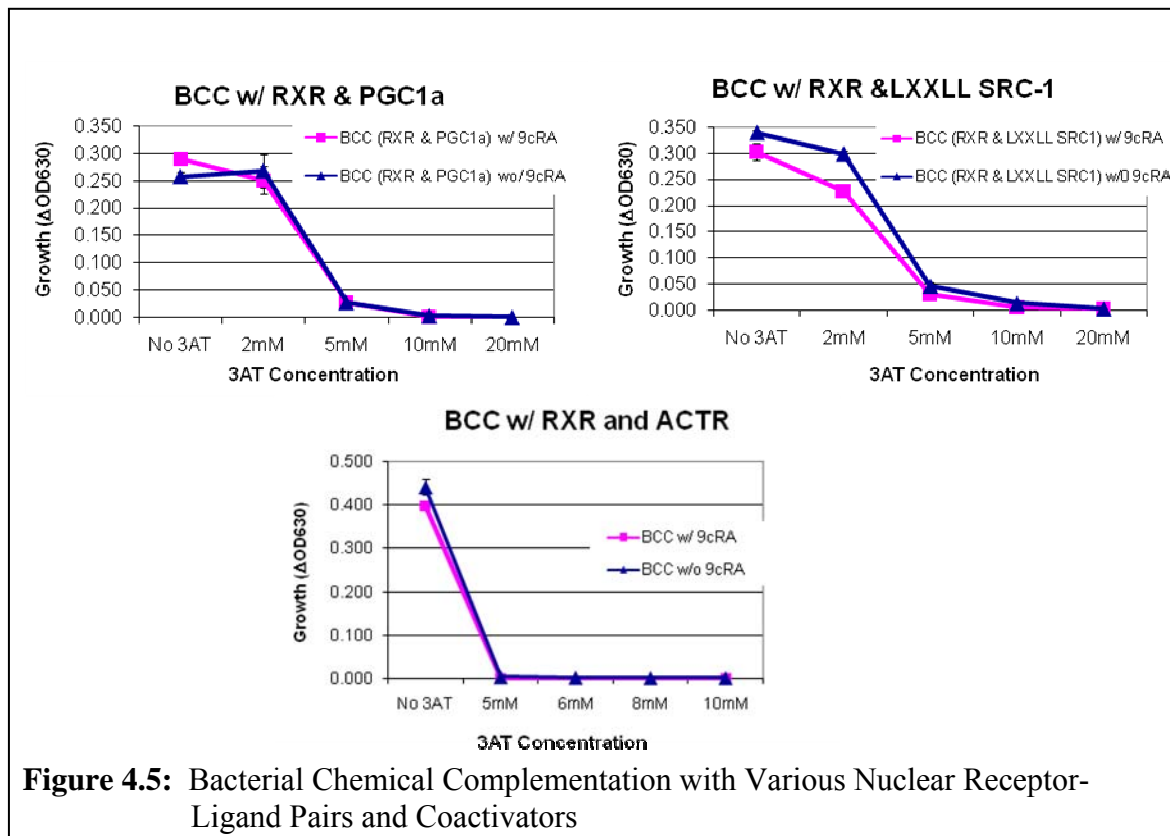
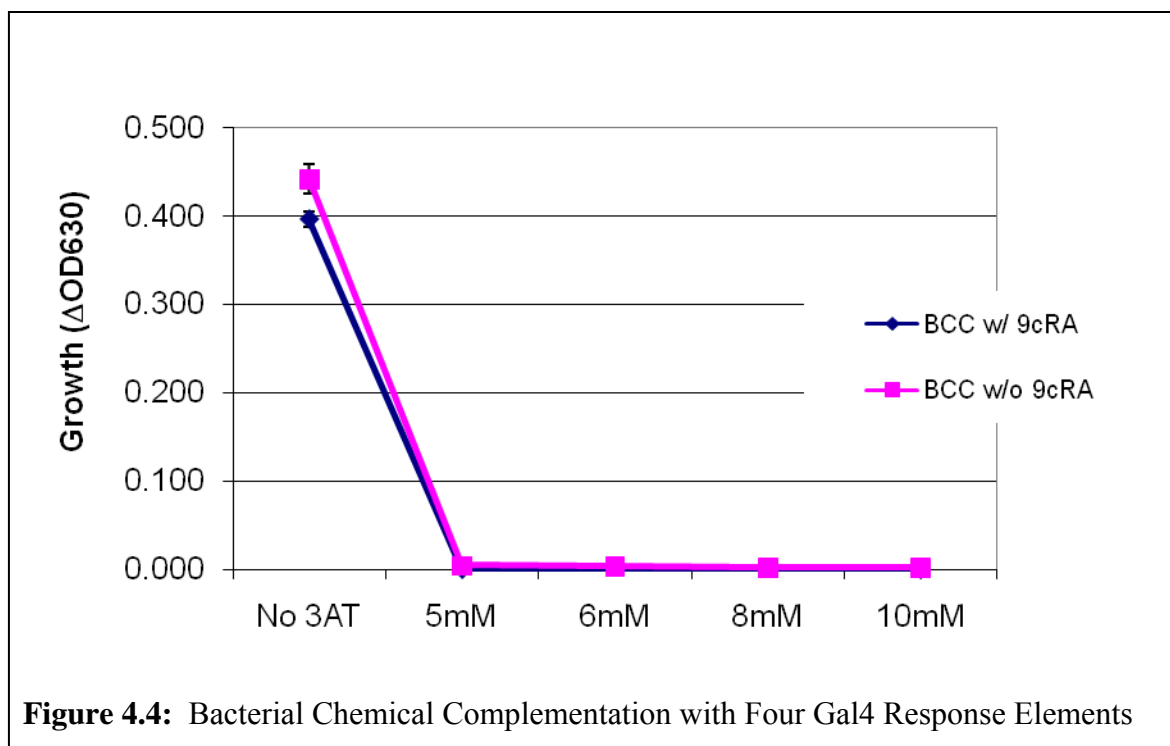
Figure 4.2: Initial Results of Ligand Independent Control

in place of the Zif268 DBDs (Figure 4.1). When the ligand independent control system was introduced into the KJIC F' Gal1p 2RE strain, containing two Gal4 consensus sequences, and tested in histidine selective media, no transcriptional activation occurred (Figure 4.2).

After further evaluating the lack of growth obtained with the ligand independent control system, previous work has shown that in mammalian cells, at least four Gal4 consensus sequences are required for a functional Gal4 based system [1, 2]. CC also uses four Gal4 consensus sequences in the Gal4 RE. Accordingly, we re-constructed the vector, pF11-Gal1p-HIS3-aadA, containing the Gal1p promoter with four response elements instead of two. After integrating this reporter construct into KJIC to create a strain with four Gal4 REs, KJIC F' Gal1p 4RE, the ligand independent control system was re-tested in histidine selective media. The KJIC F' Gal1p 4RE strain, containing four response elements, required 10 mM 3AT to remove background growth as opposed to the 5 mM needed for two Gal4 REs (Figure 4.3). The higher 3AT concentration is more than likely due to the increase in the number of response elements.

At 10 mM 3AT, growth above background was observed in our control system when four RE's were used, with an increase in OD₆₃₀ from 0.045 to 0.200 (Figure 4.3). This confirmed that our system requires the use of four REs to be functional, meaning capable of producing ligand-independent growth. Nevertheless, when testing the revised BCC system with 9cRA, the ligand for the RXR nuclear receptor, ligand-activated growth still did not occur (Figure 4.4). Although we were able to optimize the interaction between the Gal4 DBD and its response elements, these results reveal that the interaction





between the Gal4 DBD and Gal4 REs were not the sole cause of the lack of ligand-activated growth.

4.1.2 Optimizing the Various Components of the BCC system

The BCC system was designed with the nuclear receptor RXR, the ACTR coactivator, and 9cRA, the ligand for RXR. To determine if another coactivator, such as the steroid receptor coactivator-1 (SRC1) or the PPAR γ coactivator-1 (PGC-1 α), or another nuclear receptor-ligand pair were more optimal and capable of inducing ligand activated growth in our BCC system we constructed BCC systems using these various nuclear receptors and coactivators.

Additional coactivators of RXR were tested; specifically SRC1 and PGC1 α in addition to ACTR. PGC1 α is known to be a strong activator of mitochondrial function and resultantly are dominant regulators of oxidative metabolism in many tissues, specifically in tissues with high oxidative capacity, such as the heart, slow-twitch skeletal muscle, and brown adipose tissue [3, 4]. PGC1 α is a coactivator for a variety of nuclear receptors including the peroxisome proliferator-activated receptors (PPARs), retinoid receptors, estrogen receptors, farnesyl X receptor (FXR), pregnane X receptor (PXR), and liver X receptors (LXR) to name a few [5-9].

SRC1, also referred to as NCoA1, is a member of the p160/SRC coactivator family [10]. The function of the p160/SRC coactivator family members is to recruit coactivators with histone acetyltransferase activity (HAT), which will in turn assist in transcriptional activation [10-13]. SRC1, like PGC1, also activates a variety of NRs from the steroid receptors, such as the estrogen receptor, and retinoid receptors, to the various nuclear receptors which form heterodimers with RXR [14, 15]. ACTR, also referred to

as SRC-3/p/CIP/AIB1/RAC3/and TRAM-1, a member of the p160/SRC coactivator family, is involved in normal animal development to include puberty, female reproductive function, mammary gland development, and is often overexpressed in many cancers, to include breast and prostate cancer [16-18].

Both SRC1 and ACTR possess weak intrinsic HAT activities at the C-terminal region and have the ability to interact with the same nuclear receptors. Nevertheless several investigations suggest that these coactivators also have specific functions as well [19]. Specifically, ACTR exhibits greater promiscuity for various activators than the other members of the p160/SRC family, to include SRC1. Interestingly, the majority of NRs activated by SRC1 are also activated by ACTR but with different affinities [20, 21].

The hypothesis that we were not obtaining ligand activated growth due to a less than optimal interaction between RXR and ACTR could not be concluded from the BCC results obtained when testing RXR α with the 3 coactivators described above. We discovered that all three coactivators gave similar results; there was no ligand activated growth observed beyond background (Figure 4.5). To further investigate this hypothesis we tested NR-ligand pairs, with varying EC₅₀ values and fold activations, with various coactivators used.

The following BCC systems were tested for their ability to induce ligand activated growth; the pregnane X receptor (PXR) and both ACTR and SRC1 with the ligand tocopherol (7 fold activation in cell culture [22]), the liver X receptor (LXR) and SRC1 with the synthetic agonist T0901317 (T090) (EC₅₀ of 20 nM [23]), and the estrogen receptor alpha (ER α) with SRC1 with estradiol (EC₅₀ of 0.3 nM [24]). PXR is primarily involved in the detoxification of foreign toxic substances and clearing them from the

body by regulating the cytochrome p450 gene, CYP3A4, which encodes an enzyme responsible for drug metabolism. PXR, a type II NRs which heterodimerizes with RXR, is known to have a very promiscuous ligand binding pocket and as a result binds a large number of ligands to include rifampicin, paclitaxel, tocopherol, and various steroids and antibiotics, such as triacetyloleandomycin, used to treat pneumonia (Figure 4.6A). Tocopherol is a form vitamin E known to bind and activate PXR with fold activations comparable to rifampicin [22].

LXR is an important regulator of cholesterol, glucose, and fatty acid homeostasis. There are two identified isoforms of LXR, α and β , which are expressed in many of the same tissues, however, the distribution of each of the isoforms within the tissues is quite different. LXR β is expressed in almost all tissues and organs, while LXR α is restricted to the liver (where it is expressed the highest), kidney, intestine, lung macrophages, spleen and fat tissue. LXR, like PXR, activates as a heterodimers with RXR. However, this receptor is a permissive nuclear receptor in that it can be activated by 9cRA as well as its own ligands [25]. In general LXR is known to be activated by oxysterols and various synthetic ligands, to include T0901317 (Figure 4.6B). T0901317 is a synthetic nonsteroid based ligand containing multiple trifluoro functional groups that is commonly used to activate LXR in experimental studies [26].

The estrogen receptor (which is explained more extensively in Chapter 6.1) is responsible for regulation of growth, differentiation, and function in many tissues to include reproductive tracts, the cardiovascular system and mammary glands. This receptor has two forms, α and β , which are activated endogenously by the hormone 17 β -estradiol (Figure 4.6C). However, the ER, unlike LXR and PXR, are homodimers and

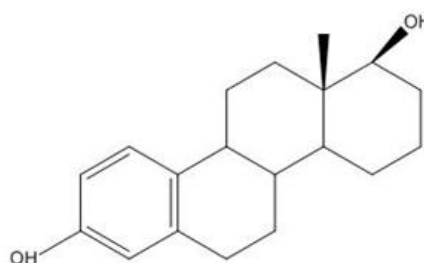
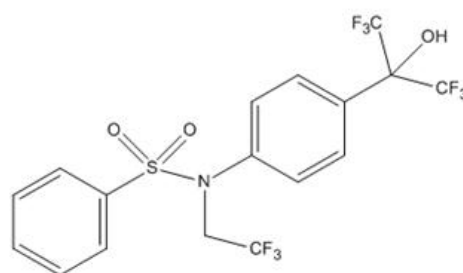
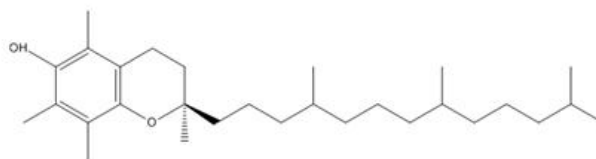
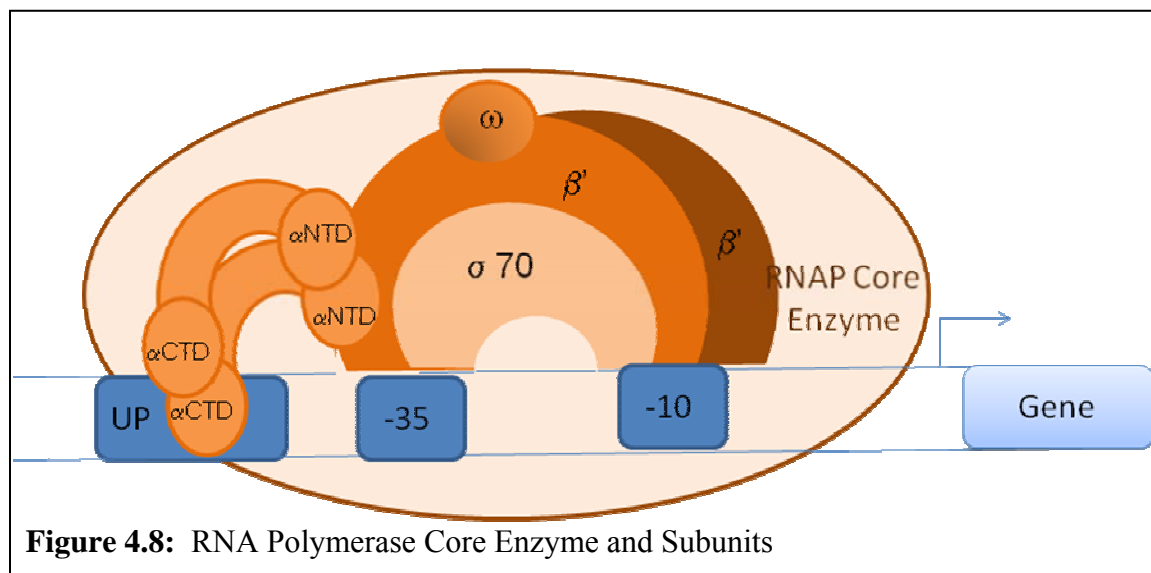
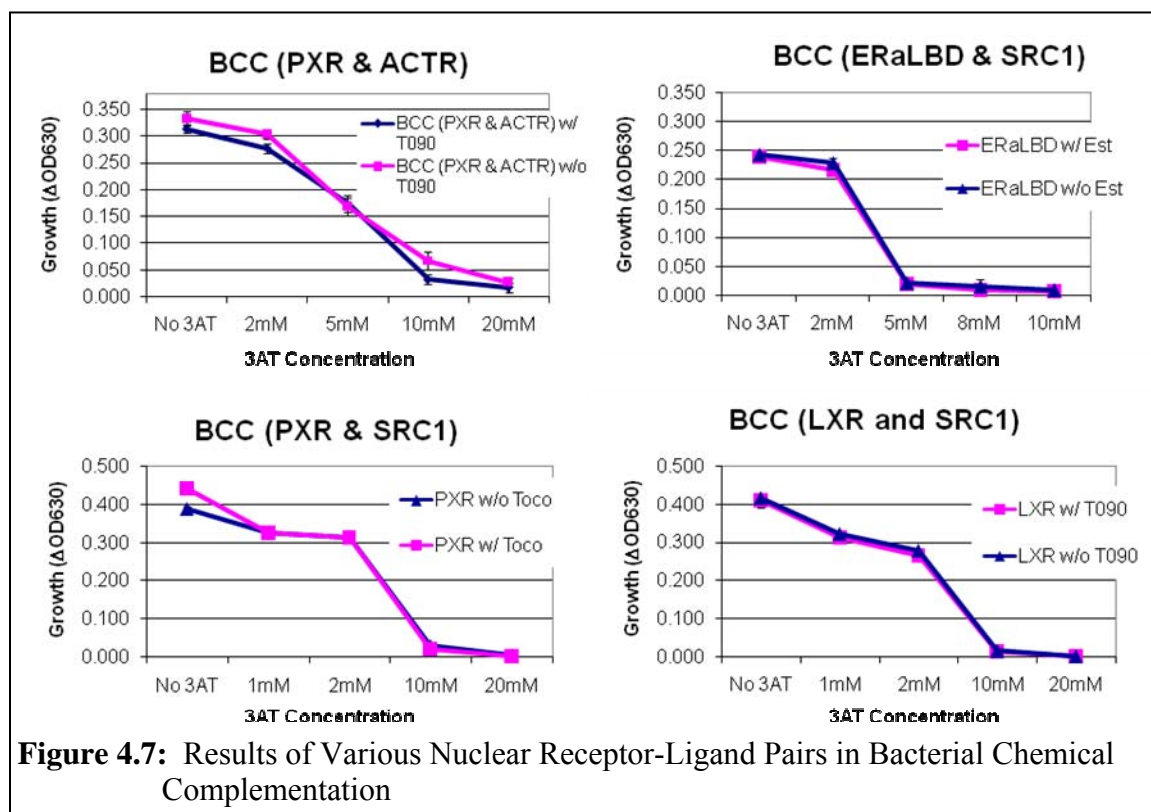


Figure 4.6: Nuclear Receptor-Ligand Pairs

are found in a multitude of tissues, to include breast, ovarian, bone, and prostate. Like other steroid receptors, ERs, unlike LXR and PXR which are found predominately in the nucleus, are cytosolic in the absence of ligand, and once ligand is bound translocate to the nucleus [27].

After creating the BCC systems with the NRs and the coactivators described above, all systems failed to produce ligand activated growth above background (Figure 4.7). With the sustained inability to obtain ligand activated growth with the BCC system, we investigated one other component of the system, the alpha subunit of RNAP. The RNAP core enzyme in *E. coli* is made up of six subunits; $\alpha 1$, $\alpha 2$, β , β' , ω , and σ (Figure 4.8). The alpha subunits both assemble the enzyme as well as recognize regulatory factors. Each subunit has a C-terminal domain (α CTD), which binds an upstream element (-40 to -70) of the promoter, and an N-terminal domain (α NTD), which binds the rest of the polymerase core enzyme. The β subunit has the polymerase activity responsible for transcription initiation and elongation. β' associates with the β subunit and binds DNA nonspecifically. The ω subunit had been known to have no discernable function for awhile but is since been found to promote assembly, however is not required for transcription *in vivo* or *in vitro* [28, 29]. Lastly the sigma subunit of RNAP is responsible for enabling specific binding of the RNAP core enzyme to the promoter regions (-35 and -10 regions). *E. coli* have seven sigma factors and each RNAP core enzyme has a specific sigma factor. Different sigma factors are activated in response to different environmental conditions. The sigma70 ($\sigma 70$) subunit is the principle sigma factor which transcribes most genes in growing bacterial cells.

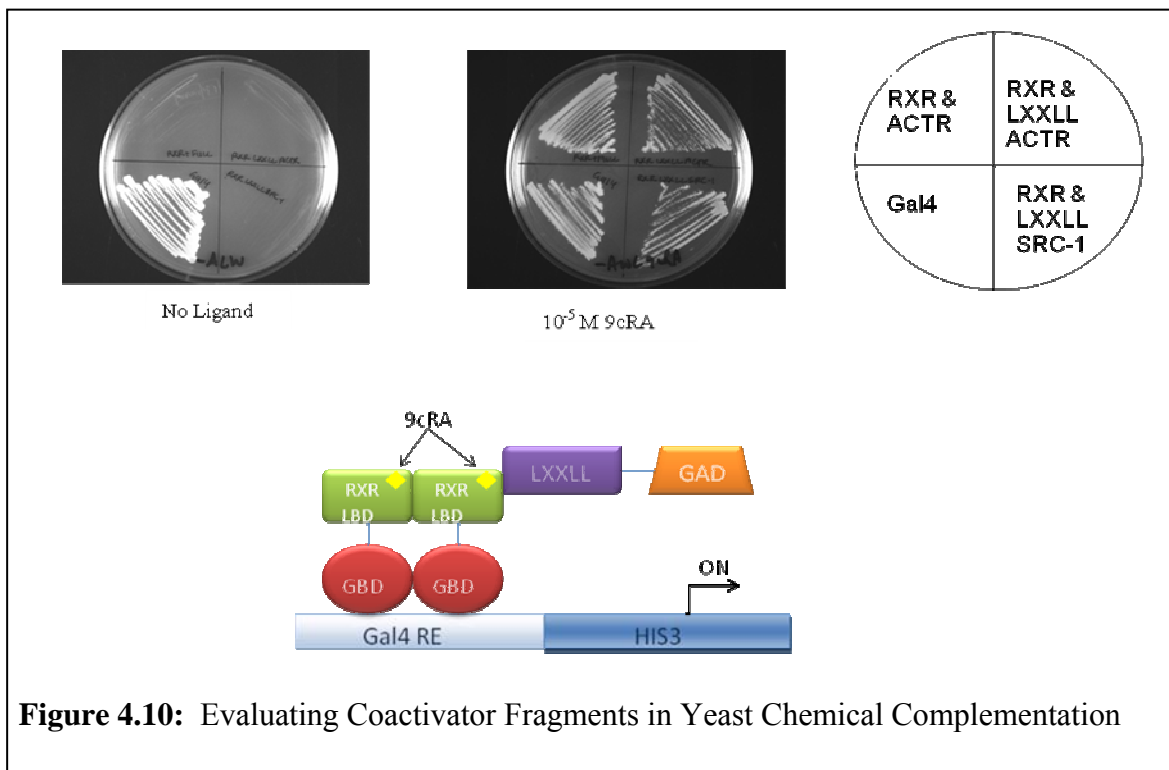
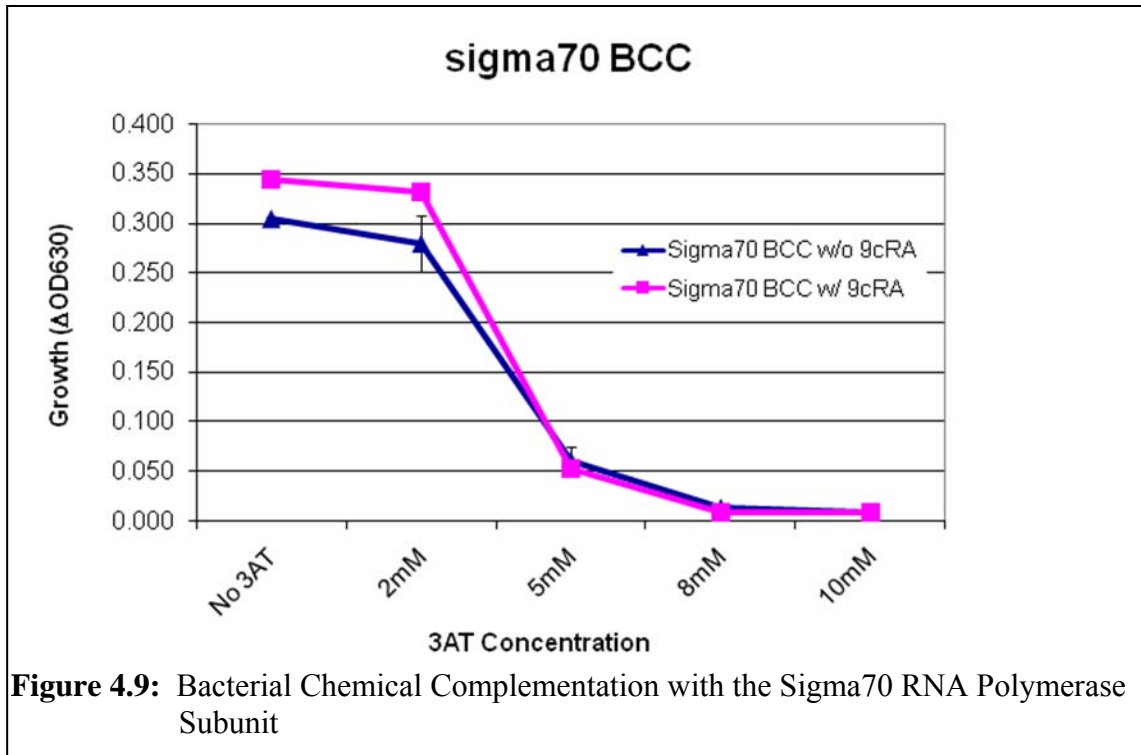


The BCC system described to this point encompasses the alpha subunit of RNAP. This subunit has been used commonly in two-hybrid systems [30, 31], along with the σ and ω subunits of RNAP [32, 33]. Both the ω and σ subunits have been tested in two-hybrid systems and were found to be successful in initiating transcription at various levels in comparison to the α RNAP subunit [32, 33]. To determine if another subunit of RNAP will have a greater affect on obtaining ligand-activated growth with the BCC system, we created a BCC system in which the coactivator protein was fused to the $\sigma 70$ subunit of RNAP. The BCC system constructed using $\sigma 70$ subunit of RNAP in place of the alpha subunit proved to not have an effect on ligand activated growth, with systems both with and without ligand having a 0.05 OD after reducing background with 5 mM 3AT (Figure 4.9).

In an attempt to obtain ligand activated growth, we tested the interactions of various components of the BCC system to see if they were causing our inability to obtain ligand dependent activation. These interactions included the GBD-Gal1p, NR-coactivator, and RNAP subunit interactions. Another avenue would be to test if the ligand is able to pass through the cellular membrane in order to bind and activate the reporter gene. However since studies in yeast have confirmed that the ligands are capable of crossing their cell membranes to enter the cell [34] and the *E. coli* cell envelope is known to be more permeable to small molecules than yeast [35], we assumed that the ligands were able to enter into the bacterial cells as well.

4.2 Enhancing Heterologous Protein Expression and Stability

After attempting to optimize the various components of the BCC systems and finding that none of the changes were capable of inducing ligand activated growth we



began to hypothesize that other factors may be contributing to the lack of ligand induced activation. Although *E. coli* are generally easy to genetically manipulate these cells do have many disadvantages as a heterologous protein expression host. Generally expression of recombinant proteins in *E. coli* can result in insoluble aggregates, known as inclusion bodies, of improperly folded and inactive proteins [36]. This occurs commonly with eukaryotic proteins which are particularly larger proteins that are post translationally modified, contain disulfide bonds, and/or are multimeric [36].

Many nuclear receptors have been found to be difficult to express in bacteria, including the glucocorticoid receptor (GR), ER, FXR, and RXR [37-40]. Due to their posttranslational modifications and multimeric nature, nuclear receptors are characteristically difficult to express in *E. coli* and as a result many systems have been developed in an attempt to enhance their heterologous protein expression [37, 40]. Due to their multimeric nature, many NRs have been separated into their DBD and LBD in order to obtain recombinant expression. Once separated the DBDs give higher soluble expression. However due to the significant hydrophobic regions of the LBD, including the ligand binding pocket and coactivator interaction surfaces, as well as post translational modifications, the LBDs tend to be prone to aggregation. Additionally because of the extremely large size of the coactivator proteins used, these proteins may also have significant challenges in relation to soluble and functional recombinant expression.

The remainder of this chapter will address the hypothesis that heterologous protein expression may be contributing to our lack of ligand activated growth. We first determine whether or not we were able to obtain soluble expression with our fusion

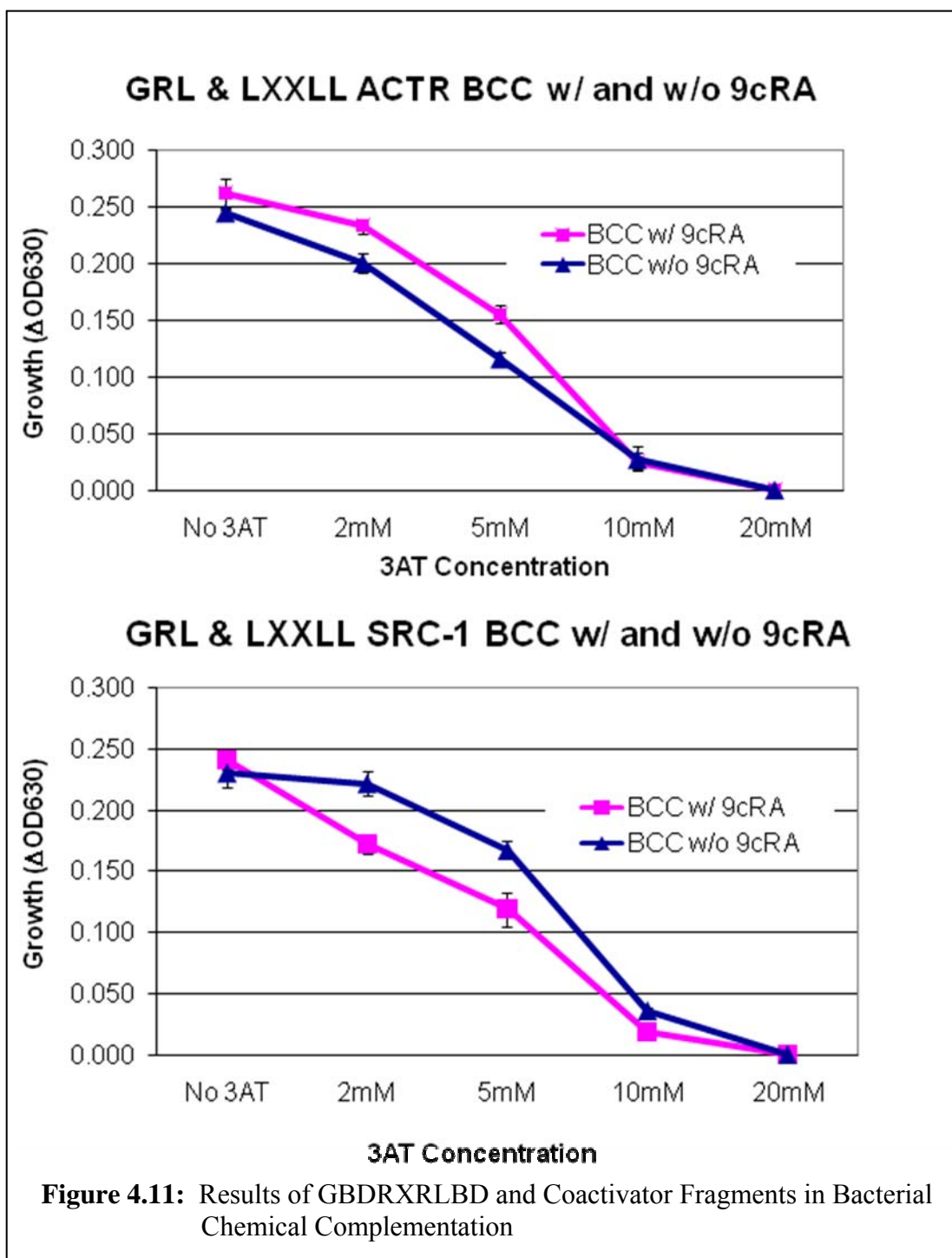
proteins in the BCC system and then we went on to attempt to enhance heterologous protein expression.

4.2.1 Analysis of Fusion Proteins Solubility

As mentioned previously, over-expression studies of nuclear receptors have shown that it is more feasible to obtain soluble protein expression when expressing the LBD alone as opposed to in conjunction with its DBD [40]. In addition some particularly larger eukaryotic proteins, such as the nuclear receptor coactivator proteins, fail to express well in *E. coli* [40-44]. Several studies reported that eukaryotic proteins are, on average, significantly larger than prokaryotic proteins [45, 46]. The full ACTR coactivator is approximately 1500 amino acids, which is an extremely large protein for bacteria to properly express. Furthermore, this coactivator has multiple post translational modifications making the protein more difficult for *E. coli* expression and folding.

Therefore, instead of using the full RXR and ACTR proteins in the BCC system, we removed the RXR DBD from the GBD-RXR fusion protein and re-constructed the α RNAP-CoA fusion proteins to contain smaller ACTR or SRC-1 fragments. These smaller, approximately 250 amino acid, fragments contain LXXLL motifs from ACTR or SCR1 that have been previously shown to interact with the nuclear receptor [47]. Since these smaller fragments had never been tested for functionality in yeast CC, a system we know to be capable of producing ligand activated growth with the correct NR-coactivator-ligand sets, we first tested these smaller coactivator fragments for function in chemical complementation in yeast.

The coactivators were cloned into yeast expression vectors as fusion with the Gal4 AD. In yeast, both the ACTR and SRC-1 fragment proteins show ligand-activated



growth with RXR and 9cRA (Figure 4.10). Unfortunately, when testing these smaller fusion proteins as well as the GBD-RXR LBD fusion protein (GRL) using BCC, again no ligand dependent activation was observed (Figure 4.11). However this data did reveal that, this new BCC system, using the GBD bound to only the LBD of RXR, required 10 mM 3AT to reduce the background (similar to the ligand independent control system), as opposed to the 5 mM required for the BCC system with the full RXR and ACTR fusion proteins. This suggests that with the full RXR BCC system used previously, the GBD-RXR construct may have had a soluble expressed GBD but an insoluble unfolded RXR protein attached to it. As a result, the GBD would still be able to bind Gal1p, but was unable to initiate transcription and as a result blocked the promoter from endogenous RNAP binding, resulting in lower background. This observation supports our hypothesis that the full RXR protein in the GBD-RXR fusion protein was not being expressed by *E. coli* and that perhaps the GRL fusion protein could increase recombinant expression.

To determine whether these newly constructed hybrid proteins were expressing in the BCC system, we over-expressed both fusion proteins under the same conditions used to test BCC. Western blot results revealed that the GRL fusion protein does appear to produce soluble protein expression. However, the α RNAP-LXXLL SRC1 fusion protein gave only insoluble expression (Figure 4.12).

4.2.2 Enhancing Heterologous Expression and Folding with Chaperones

In an attempt to enhance both folding and expression of our BCC fusion proteins in *E. coli*, we used the Chaperone Plasmid set from Takara Bio Inc to assist in protein folding. Chaperones are proteins that assist the non-covalent folding and unfolding and

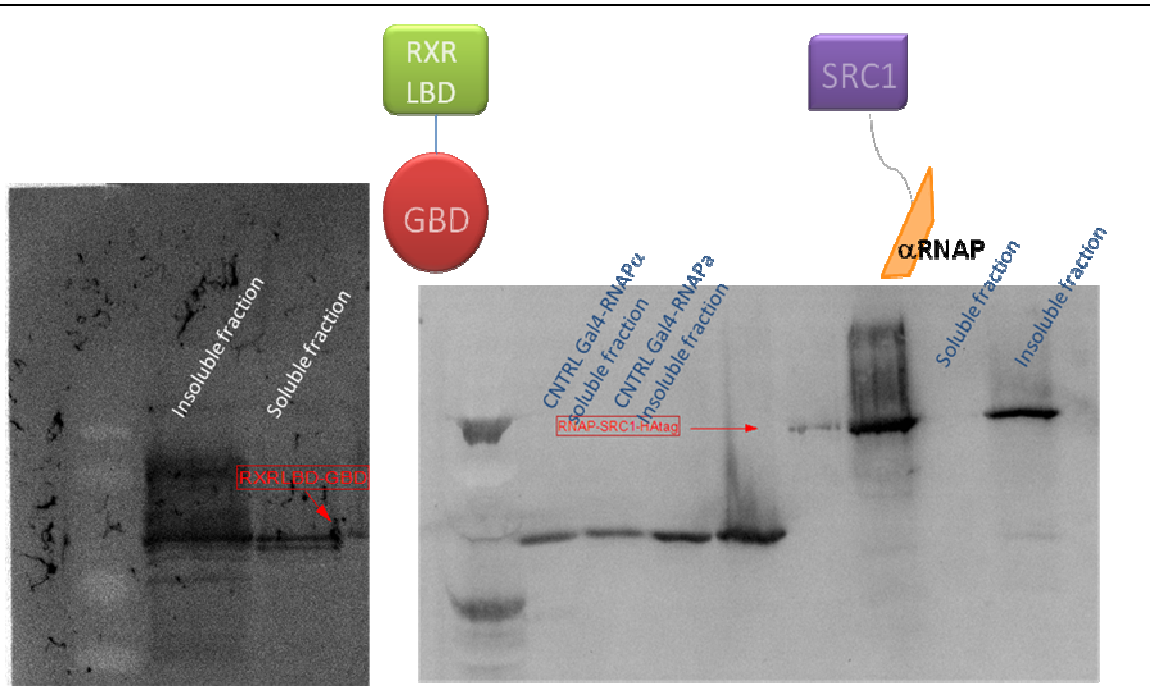


Figure 4.12: Expression of Bacterial Chemical Complementation Fusion Proteins

the assembly and disassembly of other macromolecular structures. It has previously been shown that co-expression of molecular chaperones systems with “difficult” proteins can enhance solubility [48-51]. The Takara Chaperone set contains five plasmids with different types of chaperone systems combined in an attempt to enhance protein folding and expression. These five plasmids containing various chaperone genes; the pG-KJE8 plasmid which expresses the dnaK-dnaJ-grpE and groEL-groES chaperones, pGro7 expressing only the groEL-groES chaperones, pKJE7 expression only the dnaK-dnaJ-grpE chaperones, pG-Tf2 expressing the groEL-groES-tig chaperones, and pTf16 expressing on the tig chaperones [52, 53].

There are numerous families of chaperones, each with the ability to aid protein folding in a different way [52-54]. In *E. coli* many of these proteins are highly expressed under conditions of high stress, particularly when placed in high temperatures [54]. The DnaK-DnaJ-GrpE and GroEL-GroES chaperone systems are the best characterized folding modulators in *E. coli*. The exact mechanism of how these chaperones works is not known, but they have been shown to assist in protein folding [53, 54]. GroEL/GroES may not be able to undo previous aggregation, but it does compete in the pathway of misfolding and aggregation. DnaK chaperones have been found to have a high affinity for unfolded proteins when bound to ADP, unfolded state, and a low-affinity when bound to ATP, folded state. These chaperones are thought to crowd around an unfolded substrate, stabilizing it and preventing aggregation until the unfolded molecule folds properly, at which time it lose affinity for the molecule and diffuse away (Figure 4.13).

Numerous laboratories have investigated and shown that over-expression of these chaperone proteins in conjunction with heterologous proteins in *E. coli* can facilitate

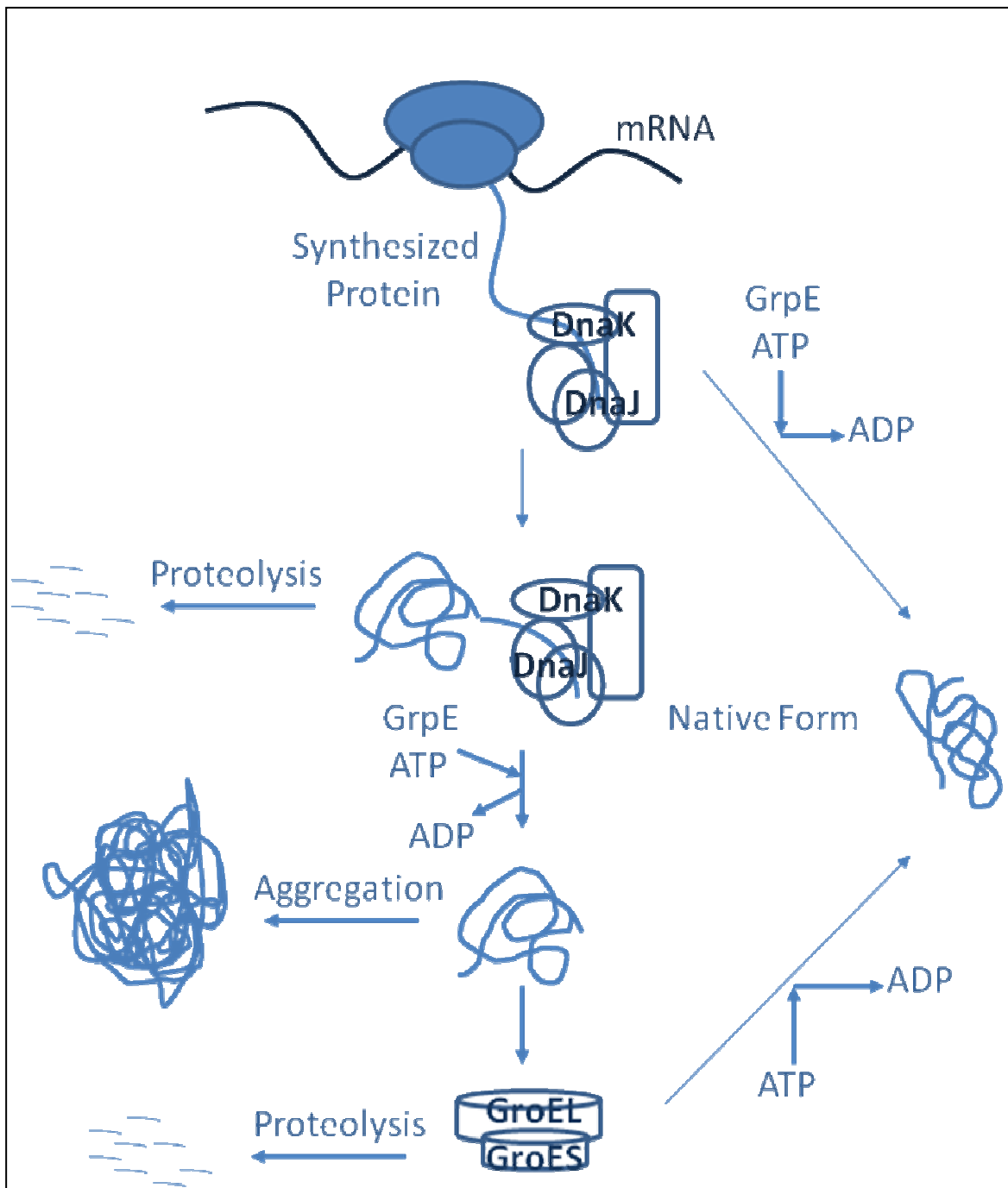
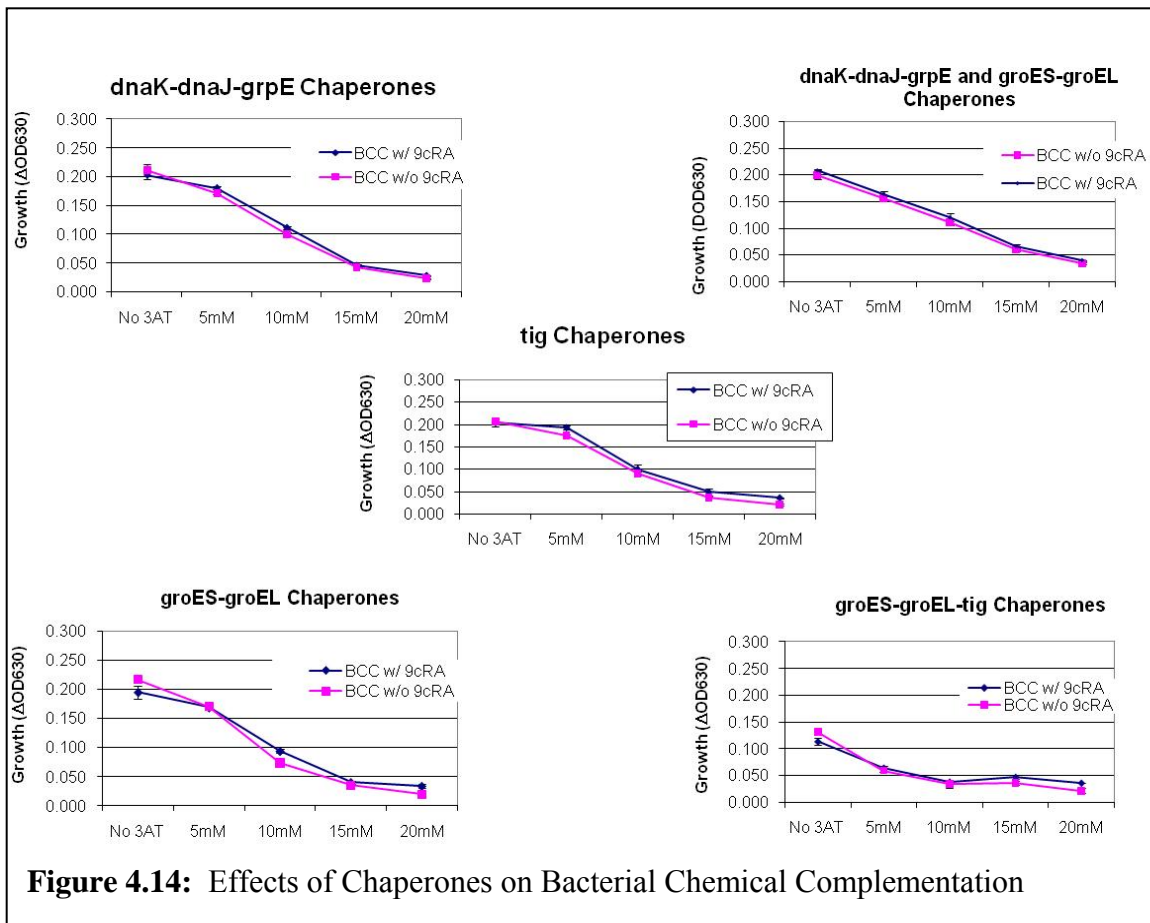


Figure 4.13: Possible Model for Chaperone Assisted Protein Folding in *E. coli*
 Revised from Nishihara, KM et al 1998 and 2000 [52, 53]

protein folding and enhance the production of functional proteins [54, 55]. The beneficial effect of higher intracellular concentrations of these chaperones, meaning folded functional recombinant proteins, is due to the facilitated folding of newly synthesized protein chains, leading to increased amounts of soluble protein vs. aggregated protein [53].

Although these chaperone systems have had great success in assisting many proteins to fold properly, there is no guarantee that chaperone co-production will improve recombinant protein expression. However, some of the recombinant proteins which have failed to express with the DnaK-DnaJ-GrpE and GroEL-GroES chaperone systems alone, are able to express when combining these proteins with additional chaperone-like factors, such as the *E. coli* trigger factor (tig) [53]. Tig is a 50 kDa protein that has been suggested to play a role in protein folding because of its association with nascent polypeptides and the 50S ribosome [56, 57]. Additionally tig has been found to associate with GroEL and strengthen its binding to proteins in order to facilitate protein folding and degradation [58].

In order to test the chaperone plasmids in with our BCC system, we first introduced the gene for the LXXLL SRC1-RNAP alpha subunit fusion protein into the *E. coli* genome. In the optimized BCC system described previously, the coactivator fusion gene is on a plasmid containing a chloramphenicol marker and the chaperone plasmids also contain chloramphenicol markers, eliminating the ability to determine which plasmid or whether both of the plasmids have been introduced into the cell. As a result, the coactivator fusion gene was conjugated into the strain along with the reporter in the same



manner as described previously to produce the strain KJIC F' Gal1p 4RE LX-S. The GRL fusion plasmid as well as each individual chaperone plasmid were then transformed into the KJIC F' Gal1p 4RE LX-S strain and tested with and without 9cRA. Unfortunately none of the chaperone systems gave ligand induced growth above background (Figure 4.14).

4.2.3 Enhancing Heterologous Expression and Folding with Osmolytes

With the advancing problem of recombinant protein expression in *E. coli*, many tools and techniques have been developed in an attempt to increase heterologous protein expression [37, 40]. One such technique to enhance protein expression is through the use of osmolytes. Osmolytes are small organic compounds which can affect the stability of proteins [59]. When cells undergo osmotic stress, resulting in a loss of water, the cell maintenance of fidelity in reactions, such as protein folding, is challenged. Reduced water activity is believed to be a critical factor for enzyme stability [60]. As a result, many cells upregulate their concentrations of specific organic small molecules, or osmolytes, to cope with the loss of water [61-63]. These osmolytes can counteract the deleterious effects of the water loss by favorably interacting with water and unfavorably interacting with the proteins backbone and side chains [64-66].

Due to their nature, osmolytes can be denaturing or protecting. Denaturing osmolytes assist in pushing proteins folding equilibrium into the unfolded states whereas protecting osmolytes push proteins into the native, or folded, state [59]. Since denatured proteins have a more exposed backbone surface than native state proteins, protecting osmolytes stabilize proteins. However since various osmolytes interact with the

backbones and side chains of proteins differently, each osmolytes impact on protein folding and stabilization varies.

Various osmolytes have been discovered and tested for their ability to control protein aggregation. Many of these organic osmolytes often accumulate in the cells cytoplasm, to include polyols, glycine, betaine, and proline. These osmolytes are considered “compatible” because they do not interfere with the cells metabolism [67]. Of these compatible osmolytes, proline has interesting properties that make it an excellent solute to use to enhance protein folding in heterologous expression systems. The first property is that proline has a remarkably high solubility in water, 7 M at ambient temperatures, making it easy to work and compatible for the cell. Additionally proline at high concentrations, greater than 1 M, has been shown to enhance the solubility of hydrophobic compounds [68].

Furthermore, proline has a distinct ability to solubilize proteins and has a more modest stabilizing effect than other osmolytes most likely due to its more favorable interaction with native state proteins side chains as opposed to its solvophobic interactions with the backbone [66, 69-71]. Because of these properties, proline studies have been done using proline at various concentrations in an attempt to promote solubility of difficult to express recombinant proteins [71-73].

Gierasch and coworkers investigated the use of proline as a chaperone for protein folding. They proposed that proline acts as a protein aggregation protector because it has the ability to suppress early aberrant protein interactions that trigger aggregation. They found that protein solubility in heterologous systems can be increased by supplementing the growth media with 300 mM of sodium chloride (NaCl), to increase the osmolality of

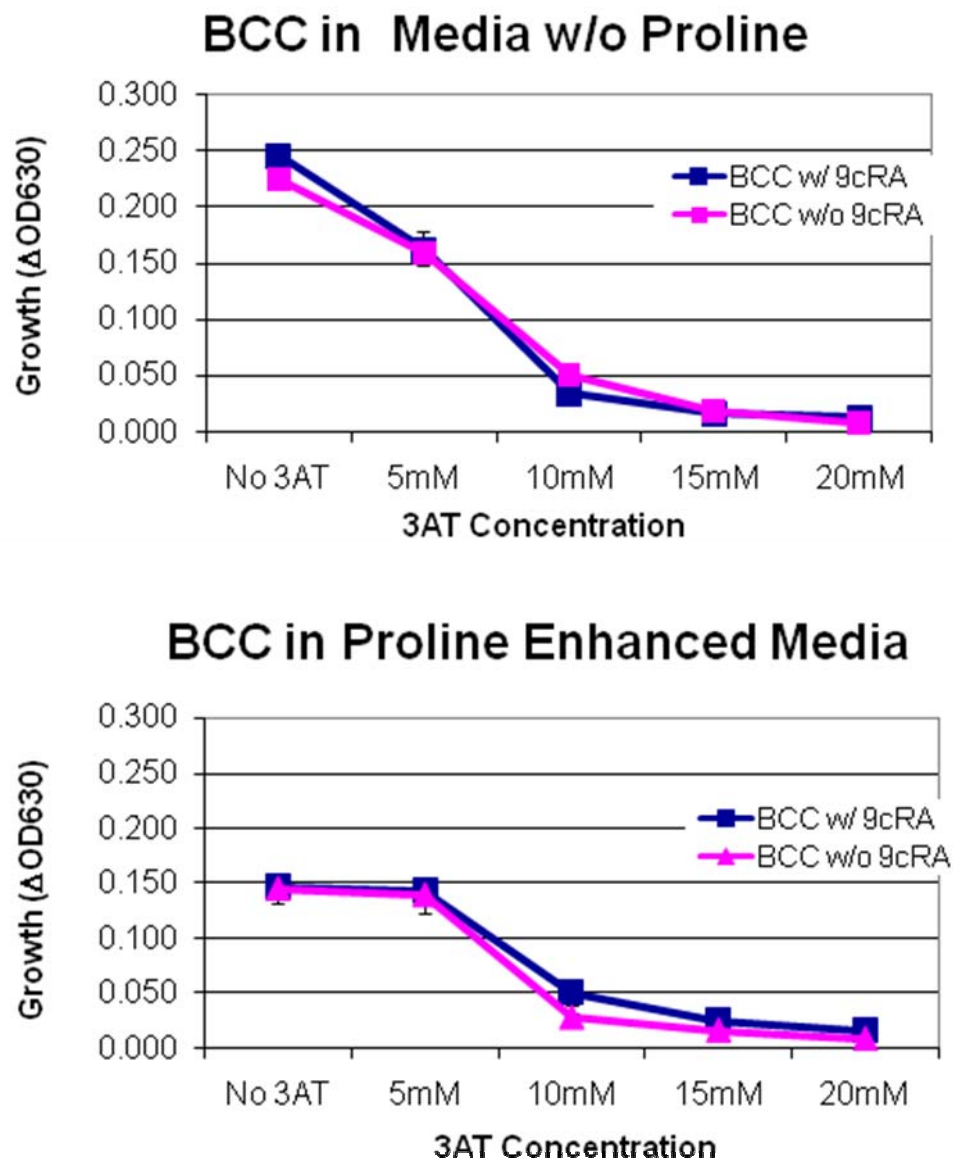


Figure 4.15: Bacterial Chemical Complementation in Proline Enhanced Media

the media, and 20 mM of proline to decrease protein aggregation and enhance soluble protein expression [72]. As a result, we tested BCC in media supplemented with 300 mM NaCl and 20 mM proline. However, the system tested with proline enhanced media reacted the same as without proline. Systems in both media required 10 μ M 3AT to reduce background. We did not see an increase in ligand induced activation above background, the OD630 was found to be 0.05 both with and without ligand. (Figure 4.15).

After further investigation into the literature, we found that some researchers found that proline as a protein folding chaperone, enhanced refolding of proteins *in vitro* when concentrations of at least 1.5 M proline were used [73]. Samuel and colleagues found that they were able to eliminate visual signs of aggregation of reduced and denatured lysozyme (r/d Lys) when adding proline at concentrations greater than 1.5 M, making proline a plausible protein folding chaperone. However, an ideal protein folding chaperone should not only prevent protein aggregation, but also restore its biological activity. Samuel and colleagues found that higher concentrations of proline not only helped to overcome protein aggregation, but also helped a fraction of the refolded protein to regain its biological activity.

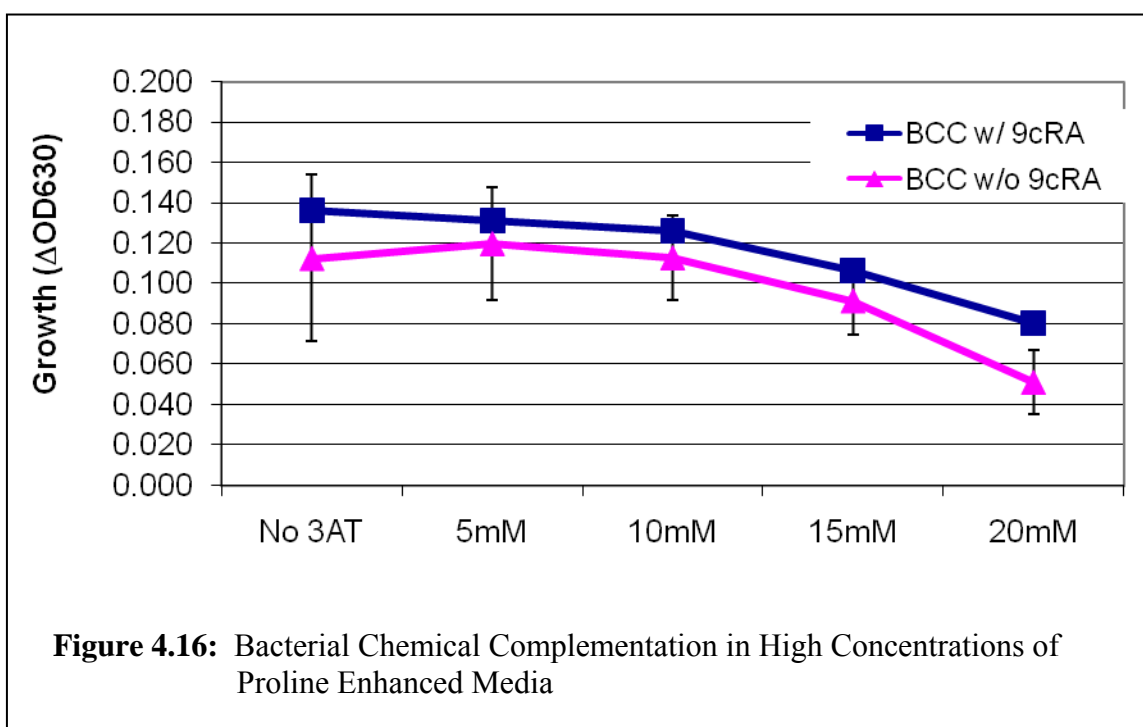
With this, we proposed to increase the proline concentration in our BCC assay media to test if we were able to enhance protein folding and function of our recombinant proteins in order to obtain ligand activated growth with our BCC system. However, the use of a higher concentration of proline *in vivo*, as opposed to *in vitro*, proved to be unsuccessful and difficult to tolerate by *E. coli*. The bacteria do not grow as well with proline concentration greater than 1.5 M. The cells grow more slowly, requiring greater

than 48 hrs as opposed to 24 hours. Additionally, these high proline concentrations cause the system to have higher background, requiring greater than 20 mM 3AT to reduce the leaky background as opposed to 10 mM previously seen using the media not supplemented with proline (Figure 4.16).

The attempts to improve heterologous protein expression of the BCC fusion proteins in *E. coli* so to enhance ligand dependent activation of BCC, we optimized the construction of the BCC system by reducing the large size of the heterologous fusion proteins, thus using the RXR LBD fused to the Gal4DBD, instead of full RXR, and smaller LXXLL based fragments of the ACTR and SRC1 coactivators, instead of the full coactivators. Additionally, we introduced chaperone proteins and/or osmolytes into the BCC system, however, none of the attempts to increase heterologous protein expression were successful, and as a result we still have a bacterial chemical complementation system unable to produce ligand activated growth.

4.3 Summary of Bacterial Chemical Complementation

Genetic selection systems have proven to be a powerful tool for evaluating macromolecular interactions, to include protein-protein, protein-DNA, and protein-small molecule interactions. After our group successfully created a yeast two-hybrid based genetic selection system to evaluate protein-small molecule interactions, referred to as chemical complementation (CC), we wanted to create an analogous system in bacteria. Both CC and BCC systems exploit the transcriptional activation function of human nuclear receptors so to link the survival of cells to the ability of these proteins to bind and activate an essential gene in response to a particular small molecule. Such a system can be used as a critical tool in drug discovery, protein and enzyme engineering, as well as



many other applications. Creating an analogous CC system in bacteria has many advantages, to include the rapid analysis of large libraries as well as the ability to evaluate small molecules not capable of being evaluated in yeast. Developing BCC has shown to produce many challenges, to include high background activity and, more importantly, inability to obtain ligand activated growth.

Although we have to yet to obtain a functional BCC system, we have been successful in creating a B2H *E. coli* strain, capable of being used to evaluate protein-protein interactions using the Gal4 DBD. Currently B2H strains contain a lambda repressor binding domain, as opposed to the GBD, to evaluate protein-protein interactions [32, 33, 74]. We have created an alternative B2H strain capable of evaluating B2H systems analogous to many Y2H systems using the gal4 DNA binding domain. The strain has currently been tested with the yeast Gal4 (dimerization domain) and Gal11p interacting proteins and found to produce activation above background with the His3 selective gene (from an OD630 of 0.05 to 0.20).

4.4 Future of Bacterial Chemical Complementation

Chemical complementation in bacteria holds great potential as a genetic selection tool for drug discovery, protein and enzyme engineering, as well as many other applications of biotechnology. However, currently the system is nonfunctional most likely due to problems with heterologous protein expression in *E. coli*. Although many attempts were made to increase recombinant protein expression, including reducing protein sizes and the use of chaperones and osmolytes, to create a functional system, a successful technique to improve heterologous protein expression must be developed. Such a technique could entail the use of other osmolytes or fusion tags, such as

thioredoxin or glutathione-S-transferase (GST). Once functional, BCC, like yeast CC, can be used in protein engineering of nuclear receptors that bind and activate in response to specific small molecules, in drug discovery, to screen through libraries of compounds that will bind and activate a particular ligand, and in deciphering biosynthetic pathways.

With the production of an alternative B2H strain, KJIC F' Gal1p 4RE, we have been able to successfully observe activation of the selectable marker using two yeast proteins previously found to interact, Gal4 and Gal11p, [32, 75, 76]. To prove that the strain is capable of evaluating mammalian proteins as well, we also intended to test the protein with two known interacting mammalian proteins,

4.5 Material and Methods

Bacterial Strains

E. coli strains, KJIC (F- Δ hisB463 Δ [gpt-proAB-arg-lac] XIII λ zaj::Tn10), and KJIC F' Z321 were a kind gift from Dr. Keith Joung (Massachusetts General Hospital/Harvard Medical School). The KJIC F' Z321 strain contains the Zif268 binding site upstream of the transcription start site of a weak *E. coli* lac-promoter and controlling the *HIS3* gene [31, 77]. The KJIC F' Gal1p 2RE and KJIC F' Gal1p 4RE strains were constructed by conjugating the Gal1p-2RE-Pwk-His3-aadA or Gal1p-4RE-Pwk-His3-aadA, respectively, construct from pF11-Gal1p-HIS3-aadA into the KJIC strain, via the CSH100 strain, using a previously described protocol [77]. The KJIC F' Gal1p 4RE LX-S strain was constructed by conjugating the Gal1p-4RE-Pwk-His3-aadA—pUV5- α RNAP-LX-S construct from pF11-Gal1p-HIS3-aadA---LX S.

Ligands

9 cis-retinoic acid (MW=300.44 g/mol) was purchased from Biomol (Plymouth Meeting, PA). Tocopherol (MW=430.69 g/mol) was purchased from Cayman chemicals (Ann Arbor, MI). T0901317 (MW= 481.3 g/mol) was purchased from Cayman chemicals (Ann Arbor, Michigan). 17- β -estradiol (MW=272.38 g/mol) was purchased MP Biomedicals Inc (Solon, OH). 10 mM stocks of the ligand were dissolved in 80% ethanol:20% DMSO (4:1 v/v) and stored at 4°C. 9cRA was stored protected from light.

Expression plasmids and Reporter Constructs

Construction of pBR-GP-Z123, pACL- α Gal4, pF11-Z321-HIS3-aadA, pBR-GP-GBDRXR, pACL- α RNAP ACTR, and pF11-Gal1p-HIS3-aadA were explained previously in Chapter 3. The chaperone plasmid set from Takara Bio Inc was a kind gift from Dr. Matsumura (Emory University, Atlanta, Ga). The pF11-Gal1p-4RE-HIS3-aadA plasmid was made exactly like the pF11-Gal1p-2RE-HIS3-aadA plasmid from Chapter 3 except with 4 Gal4 response elements instead of two.

To make pBR-GP-GBDRXRLBD, pBR-GP-GBDPXR, and pBR-GP-GBDLXR, primers containing a HindIII site at the 5' end and a AatII site at the 3' end were designed to amplify the respective fusion genes. The GBDRXRLBD fusion gene (containing residues 1-151 of Gal4 and residues 200-462 of RXR) was amplified from the pET28a-GRL vector (previously constructed in our lab). The GBDPXR and GBDhLXR α fusion genes (containing residues 1-151 of Gal4 and the full PXR or hLXR α genes) were amplified from pGBDPXR and pGBDhLXR α , respectively, (previously constructed in our lab). Once amplified, the genes were ligated into the pBR-GP-Z123-HindIII vector using the HindIII and AatII restriction enzymes.

To make pBR-GP-GBDER α LBD, the restriction site NheI was inserted into the pBR-GP-GBDRXRLBD vector at the 3' end of the GBD using site directed mutagenesis. Primers containing a NheI site at the 5' end and a AatII site at the 3' end were designed to amplify the ER α LBD fusion gene (containing residues 295-596 of ER α) from the pSG5HEGO vector and ligated into the pBR-GP-Z123-NheI vector using the NheI and AatII restriction enzymes. To make pBR-GP-Gal11p:GBD, the restriction site NcoI was inserted into the pBR-GP-Z123 vector at the 3' end of the Gal11p gene (in between the fusion of the Gal11p:Z123 fusion gene), using site directed mutagenesis. Primers containing a NcoI site at the 5' end and a AatII site at the 3' end were designed to amplify the GBD gene (containing residues 1-151 of Gal4) from the pGBDRXR vector (previously constructed in our lab) and ligated into the pBR-GP-Z123-NcoI vector using the NcoI and AatII restriction enzymes.

To make pACL- α RNAP LXXLL ACTR and pACL- α RNAP LXXLL SRC1, primers containing a NotI site at the 5' end and a AvrII site at the 3' end were designed to amplify the LXXLL ACTR fragment (containing residue 594-821 of ACTR) and the LXXLL SRC1 fragment (containing residue 594-821 of SRC1) from the pGAD10BA ACTR and pGAD10BA SRC1 vectors (previously constructed in our lab), respectively, and ligated into the pACL- α Gal4 vector using the NotI and AvrII restriction enzymes.

To make pACL- σ 70 LXXLL SRC1, the restriction site, SalI, was inserted into the pACL- α RNAP LXXLL SRC1 vector at the 5' end of the alpha subunit of RNAP, using site directed mutagenesis. Primers containing a SalI site at the 5' end and a NotI site at the 3' end were designed to amplify the Sigma70 RNAP subunit gene from the genome of the *E. coli* strain, KJIC, and ligated into the pACL- α RNAP LXXLL SRC1 -

Sall vector using the Sall and NotI restriction enzymes. To make pACL- α RNAP LXXLL SRC1-HA tag and pACL- α Gal4-HA tag, the hemagglutinin (HA) tag epitope (YPYDVPDYA) was inserted directly after the SRC1 or Gal4 genes, respectively, and before the stop via site directed mutagenesis.

To make pACL- α RNAP mPGC1, primers containing a NotI site at the 5' end and a AvrII site at the 3' end were designed to amplify the PGC1 gene (from start to stop) from the pGAD10BA mPGC1 vector (previously constructed in our lab) and ligate it into the pACL- α Gal4 vector using the NotI and AvrII restriction enzymes.

To make the yeast vectors, pGAD10BA- LXXLL ACTR and pGAD10BA- LXXLL SRC1, primers containing a BglII site at the 5' end and a NotI site at the 3' end were designed to amplify the LXXLL ACTR fragment (containing residue 594-821 of ACTR) and the LXXLL SRC1 fragment (containing residue 594-821 of SRC1) from the pGAD10BA ACTR and pGAD10BA SRC1 vectors (previously constructed in our lab), respectively, and ligated into the pGAD10BA ACTR vector using the BglII and NotI restriction enzymes.

To make pF11-Gal1p-HIS3-aadA—LX S, restriction sites PacI and AatII were inserted into the pF11-Gal1p-4RE-HIS3-aadA vector upstream of the Gal1p response element, using site directed mutagenesis. Primers containing a PacI site upstream of the pUV5 promoter and AatII site at the 3' end of the LXXLL SRC1 fragment were designed to amplify the pUV5- α RNAP-LX-S fragment from the pACL- α RNAP LXXLL SRC1 vector (previously constructed in our lab) and ligated into the pF11-Gal1p-4RE-HIS3-aadA -PacI/AatII vector using the PacI and AatII restriction enzymes.

After digestion and standard ligation, each vector was transformed into Zcompetent (Zymo Research, Orange, CA) XL1-Blue cells and selected for growth onto LB plates with the appropriate antibiotics (50 mg/ml Ampicillin, 25 µg/ml of chloramphenicol, or 30 µg/ml kanamycin, respectively). The resulting vectors were purified using QIAprep Spin Miniprep Kit (Qiagen Inc., Valencia, CA), diagnosed with restriction enzymes and sequenced for confirmation.

Bacterial Selection Protocols

All BCC systems were tested for ligand activated growth by transforming the respective expression plasmids into the respective strains using the protocol described in Chapter 3.

Media

Luria-Bertani Media (LB) media was made by dissolving 25 g of Luria broth powder in 1 L of distilled water. The pH was adjusted to 7.0 with either HCl or NaOH as appropriate and sterilize by autoclaving. M9-His selective media was made according to the BacterioMatch® II Two-Hybrid System Vector Kit (Stratagene, La Jolla, CA)) with the appropriate ligands (10 µM) and antibiotics (50 mg/ml Ampicillin (Amp), 25 µg/ml of chloramphenicol (Chl), and 30 µg/ml kanamycin (Kan)). Proline enhanced M9-His selective media was made by adding 300 mM NaCl and 20 mM proline. High proline concentrated media was made by adding 300 mM NaCl and either 1.5 M or 2 M proline. M9-His chaperone media was made by adding 4 mg/ml of L-arabinose and 10 ng/ml of tetracycline according to the Takara Bio Inc (Madison, WI). 3-amino-1, 2,4-triazole (3-AT) was purchased from Sigma (St. Louis, MO), dissolved in water to make 100 mM stock solutions, filter sterilized and stored at room temperature.

Yeast Transformations and Selective Plates

Synthetic complete (SC) plates were made as described previously [78]. Selective plates were made of SC media minus, either, leucine and tryptophan or minus adenine, leucine and tryptophan. 10 μ M 9cRA was added to the plates after the media was cooled to about 50 °C. pGBDRXR and either pGAD10BA ACTR , pGAD10BA SRC1, pGAD10BA- LXXLL ACTR, or pGAD10BA- LXXLL SRC1 were introduced into PJ69-4A using the LiAc transformation method [79]. Cells were plated onto SC –Leu-Trp plates and restreaked onto SC –Ade-Trp. Plates were incubated at 30 °C for three days.

Overexpressions and Western blots

Whole cell lysates were prepared by growing the systems in LB media with the appropriate ligands overnight. Cells were washed and re-innoculated into M9-His selective media and grown at 37°C and 300 rpm for 16-24 hrs. Cells were lysed using lysozyme and sonication. Proteins in the whole cell lysates were resolved on SDS polyacrylamide gels and transferred to nitrocellulose membranes. Western blot analyses were done using either the antibody for RXR LBD Santa Cruz Biotechnology Inc (Santa Cruz, CA) or for the HA tag , from Bethyl Laboratories Inc (Montgomery, TX). Immunodetection was done using the ECF Western Blotting Kit (Amersham Biosciences, Piscataway, NJ) according to the manufacturer's instructions.

4.6 Literature Cited

1. Emami KH & Carey M (1992) A synergistic increase in potency of a multimerized VP16 transcriptional activation domain. *Embo J* **11**, 5005-5012.
2. Sadowski I, Ma J, Triezenberg S & Ptashne M (1988) GAL4-VP16 is an unusually potent transcriptional activator. *Nature* **335**, 563-564.

3. Wu Z, Puigserver P, Andersson U, Zhang C, Adelmant G, Mootha V, Troy A, Cinti S, Lowell B, Scarpulla RC & Spiegelman BM (1999) Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1. *Cell* **98**, 115-124.
4. St-Pierre J, Lin J, Krauss S, Tarr PT, Yang R, Newgard CB & Spiegelman BM (2003) Bioenergetic analysis of peroxisome proliferator-activated receptor gamma coactivators 1alpha and 1beta (PGC-1alpha and PGC-1beta) in muscle cells. *J Biol Chem* **278**, 26597-26603.
5. Zhang Y, Castellani LW, Sinal CJ, Gonzalez FJ & Edwards PA (2004) Peroxisome proliferator-activated receptor-gamma coactivator 1alpha (PGC-1alpha) regulates triglyceride metabolism by activation of the nuclear receptor FXR. *Genes Dev* **18**, 157-169.
6. Vega RB, Huss JM & Kelly DP (2000) The coactivator PGC-1 cooperates with peroxisome proliferator-activated receptor alpha in transcriptional control of nuclear genes encoding mitochondrial fatty acid oxidation enzymes. *Mol Cell Biol* **20**, 1868-1876.
7. Wang YX, Lee CH, Tien S, Yu RT, Ham J, Kang H & Evans RM (2003) Peroxisome-proliferator-activated receptor delta activates fat metabolism to prevent obesity. *Cell* **113**, 159-170.
8. Bhalla S, Ozalp C, Fang S, Xiang L & Kemper JK (2004) Ligand-activated pregnane X receptor interferes with HNF-4 signaling by targeting a common coactivator PGC-1alpha. Functional implications in hepatic cholesterol and glucose metabolism. *J Biol Chem* **279**, 45139-45147.
9. Lin J, Yang R, Tarr PT, Wu PH, Handschin C, Li S, Yang W, Pei L, Uldry M, Tontonoz P, Newgard CB & Spiegelman BM (2005) Hyperlipidemic effects of dietary saturated fats mediated through PGC-1beta coactivation of SREBP. *Cell* **120**, 261-273.
10. Onate SA, Boonyaratanakornkit V, Spencer TE, Tsai SY, Tsai MJ, Edwards DP & O'Malley BW (1998) The steroid receptor coactivator-1 contains multiple receptor interacting and activation domains that cooperatively enhance the activation function 1 (AF1) and AF2 domains of steroid receptors. *J Biol Chem* **273**, 12101-12108.
11. Jenster G, Spencer TE, Burcin MM, Tsai SY, Tsai MJ & O'Malley BW (1997) Steroid receptor induction of gene transcription: a two-step model. *Proc Natl Acad Sci U S A* **94**, 7879-7884.
12. Bannister AJ & Kouzarides T (1996) The CBP co-activator is a histone acetyltransferase. *Nature* **384**, 641-643.
13. Ogryzko VV, Schiltz RL, Russanova V, Howard BH & Nakatani Y (1996) The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87**, 953-959.

14. Heery DM, Kalkhoven E, Hoare S & Parker MG (1997) A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **387**, 733-736.
15. Henttu PM, Kalkhoven E & Parker MG (1997) AF-2 activity and recruitment of steroid receptor coactivator 1 to the estrogen receptor depend on a lysine residue conserved in nuclear receptors. *Mol Cell Biol* **17**, 1832-1839.
16. Xu J, Liao L, Ning G, Yoshida-Komiya H, Deng C & O'Malley BW (2000) The steroid receptor coactivator SRC-3 (p/CIP/RAC3/AIB1/ACTR/TRAM-1) is required for normal growth, puberty, female reproductive function, and mammary gland development. *Proc Natl Acad Sci U S A* **97**, 6379-6384.
17. Gnanapragasam VJ, Leung HY, Pulimood AS, Neal DE & Robson CN (2001) Expression of RAC 3, a steroid hormone receptor co-activator in prostate cancer. *Br J Cancer* **85**, 1928-1936.
18. Anzick SL, Kononen J, Walker RL, Azorsa DO, Tanner MM, Guan XY, Sauter G, Kallioniemi OP, Trent JM & Meltzer PS (1997) AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer. *Science* **277**, 965-968.
19. Xu J & Li Q (2003) Review of the in vivo functions of the p160 steroid receptor coactivator family. *Mol Endocrinol* **17**, 1681-1692.
20. Bramlett KS, Wu Y & Burris TP (2001) Ligands specify coactivator nuclear receptor (NR) box affinity for estrogen receptor subtypes. *Mol Endocrinol* **15**, 909-922.
21. Heery DM, Hoare S, Hussain S, Parker MG & Sheppard H (2001) Core LXXLL motif sequences in CREB-binding protein, SRC1, and RIP140 define affinity and selectivity for steroid and retinoid receptors. *J Biol Chem* **276**, 6695-6702.
22. Landes N, Pfluger P, Kluth D, Birringer M, Ruhl R, Bol GF, Glatt H & Brigelius-Flohe R (2003) Vitamin E activates gene expression via the pregnane X receptor. *Biochem Pharmacol* **65**, 269-273.
23. Schultz JR, Tu H, Luk A, Repa JJ, Medina JC, Li L, Schwendner S, Wang S, Thoolen M, Mangelsdorf DJ, Lustig KD & Shan B (2000) Role of LXRs in control of lipogenesis. *Genes Dev* **14**, 2831-2838.
24. Kostelac D, Rechkemmer G & Briviba K (2003) Phytoestrogens modulate binding response of estrogen receptors alpha and beta to the estrogen response element. *J Agric Food Chem* **51**, 7632-7635.
25. Teupser D, Kretzschmar D, Tennert C, Burkhardt R, Wilfert W, Fengler D, Naumann R, Sippel AE & Thiery J (2008) Effect of macrophage overexpression of murine liver X receptor-alpha (LXR-alpha) on atherosclerosis in LDL-receptor deficient mice. *Arterioscler Thromb Vasc Biol* **28**, 2009-2015.

26. Wojcicka G, Jamroz-Wisniewska A, Horoszewicz K & Beltowski J (2007) Liver X receptors (LXRs). Part I: structure, function, regulation of activity, and role in lipid metabolism. *Postepy Hig Med Dosw (Online)* **61**, 736-759.
27. Htun H, Holth LT, Walker D, Davie JR & Hager GL (1999) Direct visualization of the human estrogen receptor alpha reveals a role for ligand in the nuclear distribution of the receptor. *Mol Biol Cell* **10**, 471-486.
28. Gentry D, Xiao H, Burgess R & Cashel M (1991) The omega subunit of Escherichia coli K-12 RNA polymerase is not required for stringent RNA control in vivo. *J Bacteriol* **173**, 3901-3903.
29. Heil A & Zillig W (1970) Reconstitution of bacterial DNA-dependent RNA-polymerase from isolated subunits as a tool for the elucidation of the role of the subunits in transcription. *FEBS Lett* **11**, 165-168.
30. Dove SL, Joung JK & Hochschild A (1997) Activation of prokaryotic transcription through arbitrary protein-protein contacts. *Nature* **386**, 627-630.
31. Joung JK, Ramm EI & Pabo CO (2000) A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. *Proc Natl Acad Sci U S A* **97**, 7382-7387.
32. Dove SL & Hochschild A (1998) Conversion of the omega subunit of Escherichia coli RNA polymerase into a transcriptional activator or an activation target. *Genes Dev* **12**, 745-754.
33. Dove SL & Hochschild A (2001) Bacterial two-hybrid analysis of interactions between region 4 of the sigma(70) subunit of RNA polymerase and the transcriptional regulators Rsd from Escherichia coli and AlgQ from Pseudomonas aeruginosa. *J Bacteriol* **183**, 6413-6421.
34. Schwimmer LJ (2005) Engineering ligand-receptor pairs for small molecule control of transcription. In.
35. Vaara M (1992) Agents that increase the permeability of the outer membrane. *Microbiol Rev* **56**, 395-411.
36. Weickert MJ, Doherty DH, Best EA & Olins PO (1996) Optimization of heterologous protein production in Escherichia coli. *Curr Opin Biotechnol* **7**, 494-499.
37. Cura V, Gangloff M, Eiler S, Moras D & Ruff M (2008) Cleaved thioredoxin fusion protein enables the crystallization of poorly soluble ERalpha in complex with synthetic ligands. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **64**, 54-57.
38. Schoner BE, Bramlett KS, Guo H & Burris TP (2005) Reconstitution of functional nuclear receptor proteins using high pressure refolding. *Mol Genet Metab* **85**, 318-322.

39. Ohara-Nemoto Y, Stromstedt PE, Dahlman-Wright K, Nemoto T, Gustafsson JA & Carlstedt-Duke J (1990) The steroid-binding properties of recombinant glucocorticoid receptor: a putative role for heat shock protein hsp90. *J Steroid Biochem Mol Biol* **37**, 481-490.
40. Mossakowska DE (1998) Expression of nuclear hormone receptors in *Escherichia coli*. *Curr Opin Biotechnol* **9**, 502-505.
41. Bourguet W, Ruff M, Bonnier D, Granger F, Boeglin M, Chambon P, Moras D & Gronemeyer H (1995) Purification, functional characterization, and crystallization of the ligand binding domain of the retinoid X receptor. *Protein Expr Purif* **6**, 604-608.
42. Leid M (1994) Ligand-induced alteration of the protease sensitivity of retinoid X receptor alpha. *J Biol Chem* **269**, 14175-14181.
43. Li C, Schwabe JW, Banayo E & Evans RM (1997) Coexpression of nuclear receptor partners increases their solubility and biological activities. *Proc Natl Acad Sci U S A* **94**, 2278-2283.
44. Chen ZP, Shemshedini L, Durand B, Noy N, Chambon P & Gronemeyer H (1994) Pure and functionally homogeneous recombinant retinoid X receptor. *J Biol Chem* **269**, 25770-25776.
45. Zhang J (2000) Protein-length distributions for the three domains of life. *Trends Genet* **16**, 107-109.
46. Skovgaard M, Jensen LJ, Brunak S, Ussery D & Krogh A (2001) On the total number of genes and their length distribution in complete microbial genomes. *Trends Genet* **17**, 425-428.
47. Collingwood TN, Wagner R, Matthews CH, Clifton-Bligh RJ, Gurnell M, Rajanayagam O, Agostini M, Fletterick RJ, Beck-Peccoz P, Reinhardt W, Binder G, Ranke MB, Hermus A, Hesch RD, Lazarus J, Newrick P, Parfitt V, Raggatt P, de Zegher F & Chatterjee VK (1998) A role for helix 3 of the TRbeta ligand-binding domain in coactivator recruitment identified by characterization of a third cluster of mutations in resistance to thyroid hormone. *Embo J* **17**, 4760-4770.
48. Georgiou GV, P (1996) Expression of correctly folded proteins in *Escherichia coli*. *Current Opinion in Biotechnology* **7**, 190-197.
49. Guise AW, SM; Chaudhun, JB (1996) Protein folding in vivo and renaturation of recombinant proteins from inclusion bodies. *Molecular Biotechnology* **6**, 53.
50. Hockney RC (1994) Recent developments in heterologous protein production in *Escherichia coli*. *Trends in Biotechnology* **12**, 456-463.

51. Wall JP, A (1995) Effects of overexpressing folding modulators on the in vivo folding of heterologous proteins in *Escherichia coli*. *Current Opinion in Biotechnology* **6**, 507-516.
52. Nishihara KM, K; Masanari, K; Hideki, Y; Takashi, Y (1998) Chaperone Coexpression Plasmids: Differential and Synergistic Roles of DnaK-DnaJ-GroE and GroEL-GroES in Assisting Folding of an Allergen of Japanese Cedar Pollen, Cryj2, in *E. coli*. *Applied Environmental Microbiology* **64**, 1694-1699.
53. Nishihara KM, K; Hideki, Y; Takashi, Y (2000) Overexpression of Trigger Factor Prevents Aggregation of Recombinant Proteins in *E. coli*. *Applied Environmental Microbiology* **66**, 884-889.
54. Thomas JG, Ayling A & Baneyx F (1997) Molecular chaperones, folding catalysts, and the recovery of active recombinant proteins from *E. coli*. To fold or to refold. *Appl Biochem Biotechnol* **66**, 197-238.
55. Wall JG & Pluckthun A (1995) Effects of overexpressing folding modulators on the in vivo folding of heterologous proteins in *Escherichia coli*. *Curr Opin Biotechnol* **6**, 507-516.
56. Hesterkamp T, Hauser S, Lutcke H & Bukau B (1996) *Escherichia coli* trigger factor is a prolyl isomerase that associates with nascent polypeptide chains. *Proc Natl Acad Sci U S A* **93**, 4437-4441.
57. Valent QA, Kendall DA, High S, Kusters R, Oudega B & Lührink J (1995) Early events in preprotein recognition in *E. coli*: interaction of SRP and trigger factor with nascent polypeptides. *EMBO J* **14**, 5494-5505.
58. Kandror O, Sherman M, Moerschell R & Goldberg AL (1997) Trigger factor associates with GroEL in vivo and promotes its binding to certain polypeptides. *J Biol Chem* **272**, 1730-1734.
59. Street TB, DW; Rose, GD (2006) A molecular mechanism for osmolyte-induced protein stability. *Proceedings of the National Academy of Science* **103**, 13997-14002.
60. Lin TY & Timasheff SN (1994) Why do some organisms use a urea-methylamine mixture as osmolyte? Thermodynamic compensation of urea and trimethylamine N-oxide interactions with protein. *Biochemistry* **33**, 12695-12701.
61. Yancey PH, Clark ME, Hand SC, Bowlus RD & Somero GN (1982) Living with water stress: evolution of osmolyte systems. *Science* **217**, 1214-1222.
62. Yancey PH (2005) Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *J Exp Biol* **208**, 2819-2830.

63. Record MT, Jr., Courtenay ES, Cayley DS & Guttman HJ (1998) Responses of *E. coli* to osmotic stress: large changes in amounts of cytoplasmic solutes and water. *Trends Biochem Sci* **23**, 143-148.
64. Liu XF & Bagchi MK (2004) Recruitment of distinct chromatin-modifying complexes by tamoxifen-complexed estrogen receptor at natural target gene promoters in vivo. *J Biol Chem* **279**, 15050-15058.
65. Auton M & Bolen DW (2004) Additive transfer free energies of the peptide backbone unit that are independent of the model compound and the choice of concentration scale. *Biochemistry* **43**, 1329-1342.
66. Auton M & Bolen DW (2005) Predicting the energetics of osmolyte-induced protein folding/unfolding. *Proc Natl Acad Sci U S A* **102**, 15065-15068.
67. Brown AD & Simpson JR (1972) Water relations of sugar-tolerant yeasts: the role of intracellular polyols. *J Gen Microbiol* **72**, 589-591.
68. Srinivas V & Balasubramanian D (1995) Proline Is a Protein-Compatible Hydrotrope. *Langmuir* **11**, 2830-2833.
69. Bolen DW (2004) Effects of naturally occurring osmolytes on protein stability and solubility: issues important in protein crystallization. *Methods* **34**, 312-322.
70. Bolen DW & Baskakov IV (2001) The osmophobic effect: natural selection of a thermodynamic force in protein folding. *J Mol Biol* **310**, 955-963.
71. Samuel D KT, Jayaraman G, Yang PW, Yu C (1997) Proline is a protein solubilizing solute. *Biochemistry and Molecular Biology International* **41**, 235-242.
72. Ignatova ZG, L (2006) Inhibition of protein aggregation *in vitro* and *in vivo* by a natural osmoprotectant. *Proceedings of the National Academy of Science* **103**, 13357-13361.
73. Samuel D KT, Ganesh G, Jayaraman G, Yang PW, Chang MM, Trivedi VD, Wang SL, Hwang KC, Chang DK, Yu C. (2000) Proline inhibits aggregation during protein refolding. *Protein Science* **9**, 344-352.
74. Liu Y, Buck DC & Neve KA (2008) Novel interaction of the dopamine D2 receptor and the Ca²⁺ binding protein S100B: role in D2 receptor function. *Mol Pharmacol* **74**, 371-378.
75. Hidalgo P, Ansari AZ, Schmidt P, Hare B, Simkovich N, Farrell S, Shin EJ, Ptashne M & Wagner G (2001) Recruitment of the transcriptional machinery through GAL11P: structure and interactions of the GAL4 dimerization domain. *Genes Dev* **15**, 1007-1020.

76. Farrell S, Simkovich N, Wu Y, Barberis A & Ptashne M (1996) Gene activation by recruitment of the RNA polymerase II holoenzyme. *Genes Dev* **10**, 2359-2367.
77. Whipple FW (1998) Genetic analysis of prokaryotic and eukaryotic DNA-binding proteins in *Escherichia coli*. *Nucleic Acids Res* **26**, 3700-3706.
78. Frederick M. Ausubel RB, Robert E. Kingston, David D. Moore, J. G. Seidman, John A. Smith, Kevin Struhl (1999) *Short Protocols in Molecular Biology*, 4th edn. Wiley John & Sons Inc.
79. Guthrie C (2002) Guide to Yeast Genetics and Molecular Cell Biology, Part B In *Methods in Enzymology*. Academic Press.

CHAPTER 5

LIGAND-RECEPTOR CHARGE REVERSAL VIA GENETIC ENGINEERING

5.1 Engineering Nuclear Receptors and Orthogonal Ligand Receptor Pairs

For decades, both chemical and genetic approaches have been applied to alter enzyme-substrate specificities and generate new ligand-receptor pairs [1-11]. Protein-ligand engineering is a powerful tool for manipulating and studying biological systems. Nuclear receptors are excellent protein engineering targets due to their ability to use small molecules to conditionally control genes involved in various biological processes and diseases [12-15]. Thus, nuclear receptors are commonly engineered to alter their ligand-receptor interface to respond to new “unnatural” ligands, allowing control over various biological processes [1-11]. These newly reengineered receptors are capable of regulating various processes and are used to help develop new molecular tools to manipulate and study biological systems [1-11]. In the creation of new receptors with unique small molecule selectivities, it is important to also create functionally orthogonal ligand receptor pairs (OLRPs) [16]. Functionally OLRPs are those in which the mutant receptor is not activated by the natural ligand (or may have some affinity for the natural ligand but at concentrations higher than the physiological concentration [9]). Additionally functionally orthogonal ligands should have an adequately higher affinity for the reengineered receptor over the natural one [1, 6, 11] (Figure 5.1). This is important because the ability for these OLRPs to function independently of the endogenous system allows one to specifically control only the genes

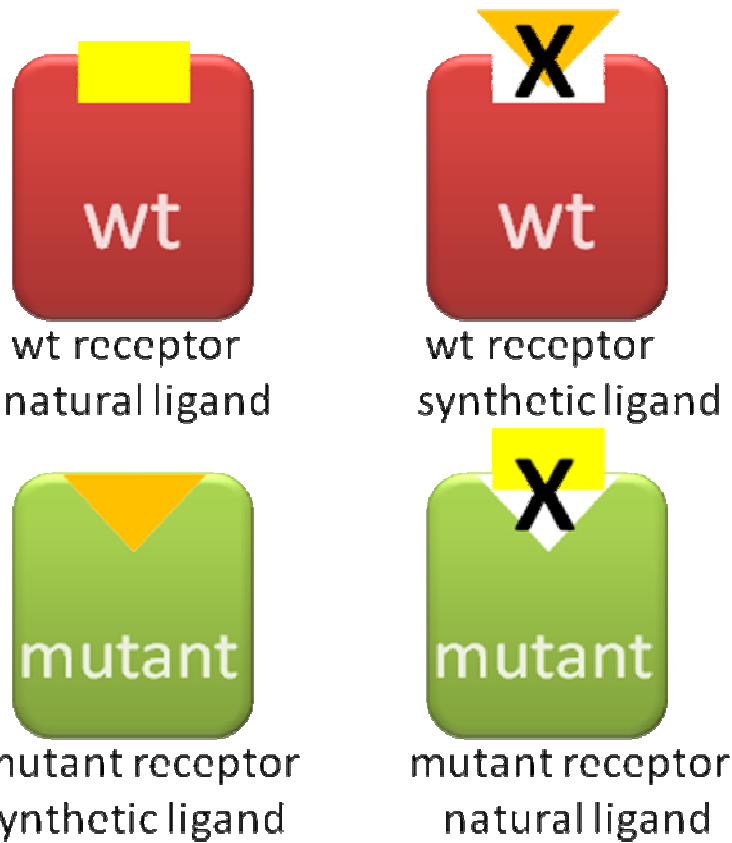


Figure 5.1: Orthogonal Ligand-Receptor Pairs

of interest.

Many approaches have been used to create these orthogonal ligand receptor pairs with varying degrees of selectivity (reviewed in [9]) including directed evolution (via random mutagenesis) as well as rational design (via site directed mutagenesis). Rational design of proteins based on structure-function relationships between ligands and receptors has gained increasing importance because of its potential to provide a general method to engineer these ligand-receptor pairs leading to a better understanding of protein structure and function [17, 18]. The ability to rationally design a specific protein for a specific substrate has obvious applications in medicine and agriculture, however doing so is an enormous challenge due to the complexity and poorly understood folding and structural aspects of proteins [19].

Nonetheless, various methods using protein engineering have been developed to create such receptors. One of the most popular methods is to alter the size of the ligand binding pockets, otherwise referred to as steric complementation strategy or the bump and hole method, and modifications of this method [20]. This approach involves mutating a large amino acid in the binding pocket of the receptor to create a “hole”, which is complemented by the addition of a functional group or “bump” on the ligand [20]. One major downfall of using bumps and holes is that the modified receptors often retain a significant affinity for the wild-type ligand [20]. Although not all applications require that the newly engineered receptor not be able to bind the wild-type ligand, these receptors cannot be used in applications which require OLRPs. The majority of

reengineered ligand-receptor pairs alter non-polar interactions, which generally result in modified receptors still having an affinity for the natural ligand [16].

With the advancements in molecular modeling and x-ray crystallography as well as protein folding and molecular recognition studies [21, 22] the presence of polar interactions have been shown to play key roles in ligand-receptor selectivity. Previous researchers have used these interactions to engineer new ligand-receptor pairs by reversing hydrogen bonding patterns [7, 11] and manipulated charged interactions, including ion pair neutralization, introducing new ion pairs, polar group exchange, and ion pair or charge reversal of various proteins (reviewed in [9]). As a result, OLRP reengineering of polar interactions has enabled discrimination between natural and engineered “unnatural” ligands [8, 11, 23, 24]. Hwang and colleagues and Tedesco and colleagues, manipulated the hydrogen bonding interaction between the ligand and receptor to create OLRPs with GTPases and estrogen receptor alpha, respectively, to accept new ligands and discriminate against their natural ligands [7, 25]. Peet and coworkers have also altered hydrophobic interactions in the binding pocket of the nuclear receptor, retinoid X receptor, to create a mutant that does not activate by wild-type ligand, 9cRA, but does activate by the synthetic compound LG335 [5].

Koh and colleagues have engineered OLRPs using a rational design strategy, called “polar exchange” [8-11, 18]. This involves replacing intra-molecular salt bridges with inter-molecular salt bridges between the ligand and receptor. Thus, a carboxylate group was removed from the Glu353 residue of estrogen receptor and replacing it with a carboxyl group on the ligand. This interaction retains a network of polar interactions similar to the native complex and differs in the covalent connectivity of key polar

functional groups [11]. In addition to polar exchange, the Koh group has also attempted to use charge reversal techniques to create OLRPs. A S298D mutation in the retinoic acid receptor gamma (RAR γ) was made to change the specificity of the ligand from the negatively charged all *trans*-retinoic acid (atRA) to a positively charged retinol derivative [10]. However, the transcriptional activation by this variant was greatly reduced in comparison to wild-type RAR- atRA activation [10]. Charge reversal could potentially be an excellent method for protein engineering to develop OLRPs. Charge reversal can exploit orthogonality in that the opposite charges engineered into the receptor makes binding to the natural ligand unfavorable.

5.2 Motivation for Engineering Ligand-Receptor Charge Reversal

The ability to rationally design a specific protein for a specific substrate has obvious applications in medicine and agriculture, including drug discovery and the production of genetically modified foods that resist pest and bacterial infection [26]. The complexity of proteins makes their engineering challenging, specifically due to the complex effects that combinations of binding forces, including electrostatic, steric, and hydrophobic effects, have on proteins structure and function [19]. Nonetheless, protein engineering of NRs can not only help to produce OLRPs for applications in medicine and agriculture, but also leads to a better understanding of protein function and structure because the mutational analysis helps to investigate the stability and tolerability of mutations. By mutating the residues within the ligand binding pocket of nuclear receptors, we can investigate changes in both function and structure. As mentioned previously, polar interactions play key roles in ligand-receptor selectivity and

reengineering of polar interactions, specifically via a charge reversal method, can enabled discrimination between natural and engineered “unnatural” ligands.

Previous attempts at protein engineering using charge reversal include that of Kirsch and coworkers, who engineered an aspartate aminotransferase to accept arginine instead of its natural ligand, aspartate. However, they were unable to produce a protein-ligand pair that exhibited wild-type levels of activity [24]. Additionally, as mentioned previously, Koh and colleagues were able to engineer the retinoic acid receptor gamma, RAR γ , charge reversal mutant, (S289D), to bind a positively charged guanidine retinoid. However, this variant yielded a modest 3-fold induction of a luciferase reporter in Hela cells [10]. A drawback to reengineering ligand-receptor recognition by charge reversal, as shown in the two examples above, is that the newly engineered pairs have markedly lower potencies and selectivities in comparison to the original pair [10, 24, 27-29]. Hwang and Warshel suggest that charge reversal will never produce variants that have wild-type-like activity because these wild-type ligand-receptor interactions are stabilized by more than the amino acids in the immediate vicinity of the salt bridge. Altering only local residues leaves these other interactions intact and does not contribute additional stability to the reversed salt bridge [30]. In other words the active site environment designed by nature to stabilize the wild-type ion pair is not optimized for the reverse pair.

Prior ion pair reversals involved altering a single residue in the binding pocket known to interact with the ligand [10, 24]. According to Hwang and Warshel, the overall potential of the pocket still favored the wild-type protein-substrate pairing, resulting in new protein-substrate pairs with lower potencies than wild-type. An alternative approach to engineering OLRPs via charge reversal so to create ligand-receptor pairs with

selectivities comparable to the natural pair would be to design a favorable polarized environment around a charge reversal ligand-receptor pair as opposed to a single charge reversal residue-ligand contact. With this notion, and the information on ligand-receptor binding interactions often obtained from crystallographic structures, a favorable polarized environment around a charge reversal ligand-receptor pair can be created by changing multiple polar residues in the binding pocket that interact with and stabilize the ligand.

Estell and colleagues engineered double mutants of subtilisin that interact with peptide substrates with charges opposite of those of the native substrates [31]. Although, they changed only two residues to create a more positively charged binding cleft, their charge reversal mutant not only had an increased affinity for the negatively charged substrate over the natural positively charged substrate but additionally had a potency with the new ligand similar to that of the wild-type subtilisin and its natural substrate [31]. Due to the success of Estell and colleagues with a designed charge reversal variant, we created a designed library of retinoic acid receptor alpha (RAR α) mutants and selected for variants that bind positively charged ligands, as opposed to the natural negatively charged ligand, all *trans* retinoic acid (atRA).

RAR is a nuclear receptor that regulates embryonic development processes and is activated by retinoids like all *trans* retinoic acid, atRA (explained further in Subchapter 5.3). The crystal structure for RAR γ and atRA was solved by Renaud and coworkers [32] which not only suggested an electrostatic guidance mechanism for ligand binding but also revealed the residues, specifically polar residues, involved in the ligand binding interactions between RAR γ , and other isotypes of RAR, and atRA (which are needed to engineer charge reversal variants). Resultantly, RAR is a great model to design an

effective charge reversal protein engineering method to create OLPRs, evaluate receptor structure and function, and to evaluate whether this method is consistent with Warshel's hypothesis.

5.3 Retinoic Acid Receptor

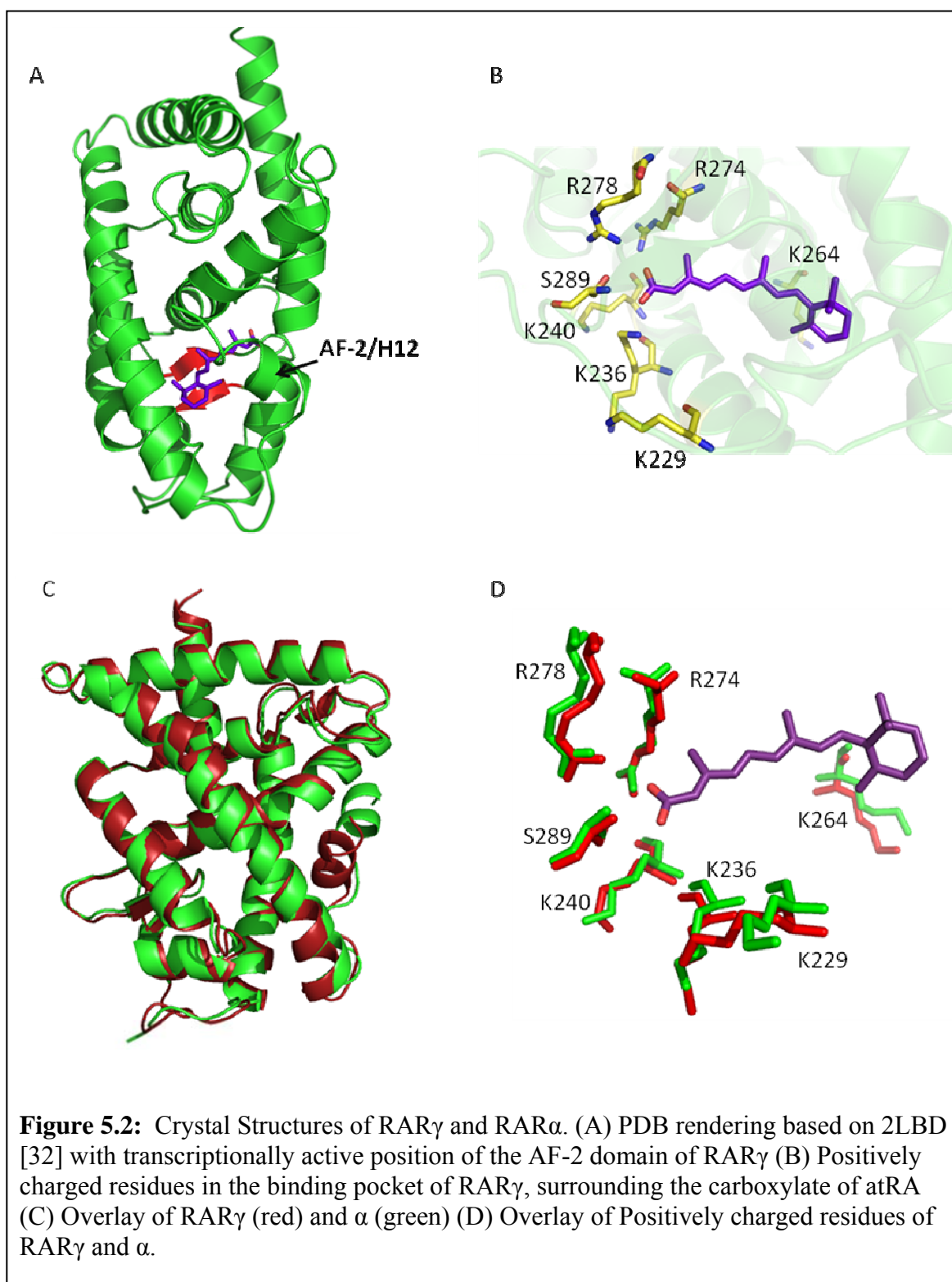
Biologically active vitamin A metabolites, also known as retinoids regulate a variety of essential biological processes, to include cell development, differentiation, homoeostasis, and apoptosis, as well as their disorders [12, 33-35]. These compounds, such as atRA and 9-*cis* retinoic acid (9cRA), exert their pleiotropic effects through two families of nuclear receptors, the retinoid X receptors (RXR) (explained in Chapter 1) and the retinoic acid receptors (RAR). RAR functions as a ligand dependent transcription factor by heterodimerizing with RXR. RAR-RXR heterodimers work in the same manner as other type II nuclear receptors (NRs) [36]. In the absence of ligand, or in presence of an antagonist, RAR-RXR heterodimers recruit corepressors, NcoR and SMRT, and histone deacetylases (HDACs), which lead to inactive and condensed chromatin, preventing transcription. Once an agonist ligand has bound, the NRs undergo a conformational change, releasing corepressors and recruiting coactivator proteins and histone acetyltransferases, leading to active expanded chromatin, and activating transcription [36-41]. Generally, RAR is considered a non-permissive RXR heterodimerization partner, in which RXRs are unable to be activated by agonists in the absence of RAR bound ligands [42].

RAR has three isoforms, α , β , and γ and is activated by both all *trans* retinoic acid and 9-*cis* retinoic acid. The DNA binding domain (DBD) as well as the ligand binding domain (LBD) of all three RAR isotypes are highly conserved within a given species,

94% - 97% and 84% – 90%, respectively. Evidence that RAR isotypes have distinct physiological functions came from the observation that RAR α is found to be present in most tissues, where as RARs β , and γ , are more selective [45].

The crystal structure for RAR γ and atRA, solved by Renaud and coworkers [32], revealed several key residues involved in the ligand binding interactions between RAR γ and atRA. One particular feature of this receptor is the electrostatic guidance mechanism for ligand binding. The elucidation of this crystal structure also revealed what Renaud called a “mouse trap” mechanism in which ligand binding induces a conformation change within the receptor that repositions the α -helix 12, also referred to as the AF-2 domain, forming a transcriptionally active receptor (Figure 5.2A). These crystallographic studies revealed that the negatively charged carboxylate of atRA is stabilized by the positive potential at the surface of RAR γ 's binding pocket. Residues K236, R278, and S289 in the binding pocket were found to anchor the carboxylate of atRA [32]. A cluster of positively charged residues, (K229, K236, K240, R274 and R278) near or at the surface of helices 3 and 5 (H3 and H5, respectively) also stabilizes the carboxylate (Figure 5.2B). K264 is at the entrance of the ligand binding cavity, and is proposed to attract atRA out of solution and guide it to the binding site. Once ligand is bound, K264 also forms salt bridges with E414 and E417 to anchor H12, the AF-2 domain, and seal the cavity [32].

The results of this crystallographic study revealed positively charged residues involved in the ligand binding interactions between RAR γ and atRA that work to stabilize this interaction. Basic amino acids, such as arginine and lysine, have been found to be involved in retinoid ligand binding in various isotypes of the retinoic acid receptors [32, 46-48]. The residues in RAR γ 's binding pocket that were found to anchor and



stabilize the carboxylate of atRA [32] are conserved in all three isotypes of the retinoic acid receptors [48]. Because of the implication of RAR α in diseases [49-51] and the similarities in structure between RAR γ and α (Figure 5.2 C & D) we chose RAR α to investigate charge reversal engineering. We created a library of RAR α charge reversal mutants based on the RAR γ residues found to be involved in producing the positive potential involved in stabilizing the natural negatively charged atRA ligand. The remainder of this chapter will focus on the library design and results of charge reversal as a protein engineering method to investigate protein structure and function, and to possibly create OLRPs.

5.4 Results of Ligand-Receptor Charge Reversal

5.4.1 Design of Engineered Receptor Library for Charge Reversal

With the high resolution crystal structure available for RAR γ , giving insight into the residues involved in ligand binding and the use of substituted side chain atoms as opposed to backbone atoms that define the binding site, we chose RAR α to design an effective charge reversal protein engineering method to evaluate protein structure and function. To create our charge reversal ligand-receptor pair, we created a “positive potential” library of RAR α variants based on the crystallographic data of RAR γ bound to atRA. This library consisted of single, double, triple, quadruple, and quintuple charge reversal mutants intended to bind positively charged ligands as opposed to the negatively charged atRA. Changing the electrostatic potential and the charge of the binding pocket used to stabilize atRA, allows for the reversal of RAR α ’s binding pocket from an overall positive environment to a negative environment possibly creating a new positively charged ligand to bind more favorably than atRA. Changing the overall charge

environment of the binding pocket, should in turn create a ligand-receptor charge reversal pair functionally comparable to RAR α -wt and atRA.

The charges of four basic residues (K229, K236, R274, and R278) and one neutral residue (S289) in the RAR α binding pocket (residue numbering based on RAR γ), were modified due to their ability to stabilize atRA directly or attribute to the positive environment involved in stabilizing the negatively charged ligand [32]. As mentioned previously, the K236, R278, and S289 residues have all been shown to be directly involved in anchoring the carboxylate of atRA. Both R278 and K236 anchor the carboxylate of atRA by forming salt bridges, while the S289 residue forms hydrogen bonds with the carboxylate. R278 not only forms a weak salt bridge with the carboxylate but also contributes to the hydrogen bonding network that anchors the carboxylate. The K229 and R274 residues chosen do not directly interact with the carboxylate of atRA, but are suggested to be a part of the cluster of positively charged residues near the surface of the binding pocket involved in the stability and electrostatic guidance of the ligand [32]. Previous results which modified the corresponding murine RAR residue K229 to an alanine, reduced atRA affinity [46], suggesting that although this residue does not make direct contact with the ligand it does stabilize the ligand by maintaining a positive potential in the binding pocket.

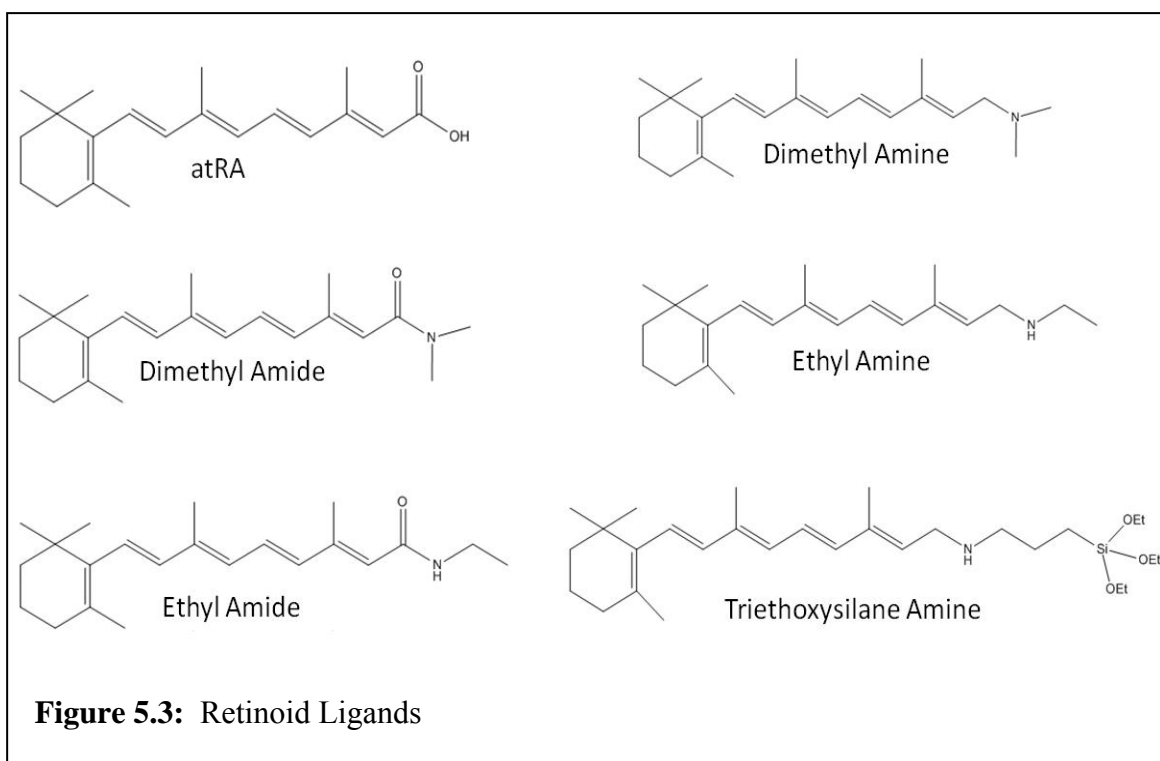
Site directed mutagenesis was used to modify the four charged residues to a glutamic acid (E) and S289 was modified to an aspartic acid (D). We hypothesized that mutating at least two or more of these sites to negatively charged aspartic acids or glutamic acids will alter the overall potential of the binding pocket to allow positively charged ligands to bind and activate RAR α . Single, double, triple, quadruple, and

quintuple mutants were constructed via site directed mutagenesis and confirmed by sequencing (Table 5.1). Five all *trans* retinoid based neutral and positively charged ligands were synthesized (courtesy of Dr. Stephan France, Georgia Institute of Technology, GA). These ligands included two amides, an ethyl amide retinoid (EthAmD RA) and a dimethyl amide retinoid (DiAmD RA), and three amines, a dimethyl amine retinoid (DiAmN RA), an ethyl amine retinoid (EthAmN RA), and a triethoxysilane amine retinoid (SiOEth RA) (Figure 5.3) [10]. Once synthesized, the variants and ligands were tested in yeast using chemical complementation to determine their activities. The hypothesis is that the positively charged amine compounds will have a decreased affinity for atRA and that the variants will have a higher affinity for the amine ligands.

As mentioned in Chapter 2, CC is a generalizable selection systems in *S. cerevisiae* used as a tool for investigating protein-small molecule interactions [52, 53]. This system uses genetic selection to link the survival of a yeast cell to the presence of a small molecule able to bind and activate RAR α . The CC system used to test the ability of the neutral or positively charged retinoids to activate RAR α or its mutants was constructed using the following two fusion proteins: The Gal4 DBD is fused to the RAR α wild-type or mutant ligand binding domains (GBD: RAR) and the human nuclear receptor co-activator, ACTR, is fused to the yeast co-activator Gal4's activation domain, GAD (ACTR:GAD). The GBD binds the Gal4 response element, regulating the transcription of the ADE2 gene. If the ligand of interest is able to bind and activate RAR α or its mutants, the LBD will undergoes a conformational change, recruiting the co-activator, ACTR:GAD, fusion protein, and initiating transcription of the ADE2 gene. Thus, the yeast cells containing a functional ligand receptor pair survive in media lacking

Table 5.1: Library of Site Directed Charge Reversal Variants

Single Mutants	Double Mutants	Triple Mutants	Quadruple Mutants	Quintuple Mutants
R274E K236E S289D R278E K229E	R274E;K236E R274E;S289D R274E;R278E R274E;K229E K236E;S289D K236E;R278E K236E;K229E S289D;R278E S289D;K229E R278E;K229E	R274E;K236E;S289D R274E;K236E;R278E R274E;K236E;K229E R274E;S289D;R278E R274E;S289D;K229E R274E;R278E;K229E K236E;S289D;R278E K236E;S289D;K229E K236E;R278E;K229E S289D;R278E;K229E	R274E;K236E;S289D;R278E R274E;K236E;S289D;K229E R274E;K236E;R278E;K229E R274E;S289D;R278E;K229E K236E;S289D;R278E;K229E	R274E;K236E;S289D;R278E;K229E



adenine (Figure 5.4).

5.4.2 Effects of Amide Based Retinoids on Charge Reversal Mutants

RAR α wild-type (RARwt) and all mutants were tested in CC with atRA, dimethyl amide (DiAmD RA), and the ethyl amide (EthAmD RA) retinoids (Table 5.2). All the single mutants, most of the double mutants (except S289D;R278E and R278E;K229E), and one of the triple mutants (R274E;K236E;K229E) are activated by the EthAmD RA as well as by the natural ligand, atRA. As shown in Table 5.2, the triple mutant R274E;K236E;K229E, activates at atRA concentrations as low as 1 μ M, and EtAmD concentrations of 10 μ M (Figure 5.5A). The results also show that both the amide ligands activate wild-type RAR α (Figure 5.5B) and none of the variant-amide ligand-receptor pairs activate as well as the RAR wt with atRA. The single variant K236E, activates at concentrations as low as 0.1 μ M with the EthAmD RA, but RAR wt activates with concentrations as low as 0.01 μ M with atRA (Figure 5.6). Despite the fact that the fold activation was not as high as RAR wt with atRA, we were able to obtain activation with many variants and the amide ligand.

All of the quadruple and quintuple mutants, along with the two double mutants, S289D;R278E and R278E;K229E, and the remainder of the triple mutants, are not activated by atRA or the amide ligands. However, all of the variants that were functional were activated equally by both atRA and EthAmD RA, with the exception of 4 variants, single mutants S289D and R278E, double mutant R274E;K236E, and triple mutant R274E;K236E;K229E, which all have higher activation with atRA than the EthAmD (Table 5.2). This inferred that generally, most of these mutations, individually or in combination, do not discriminate between the negatively charged ligand and a neutral

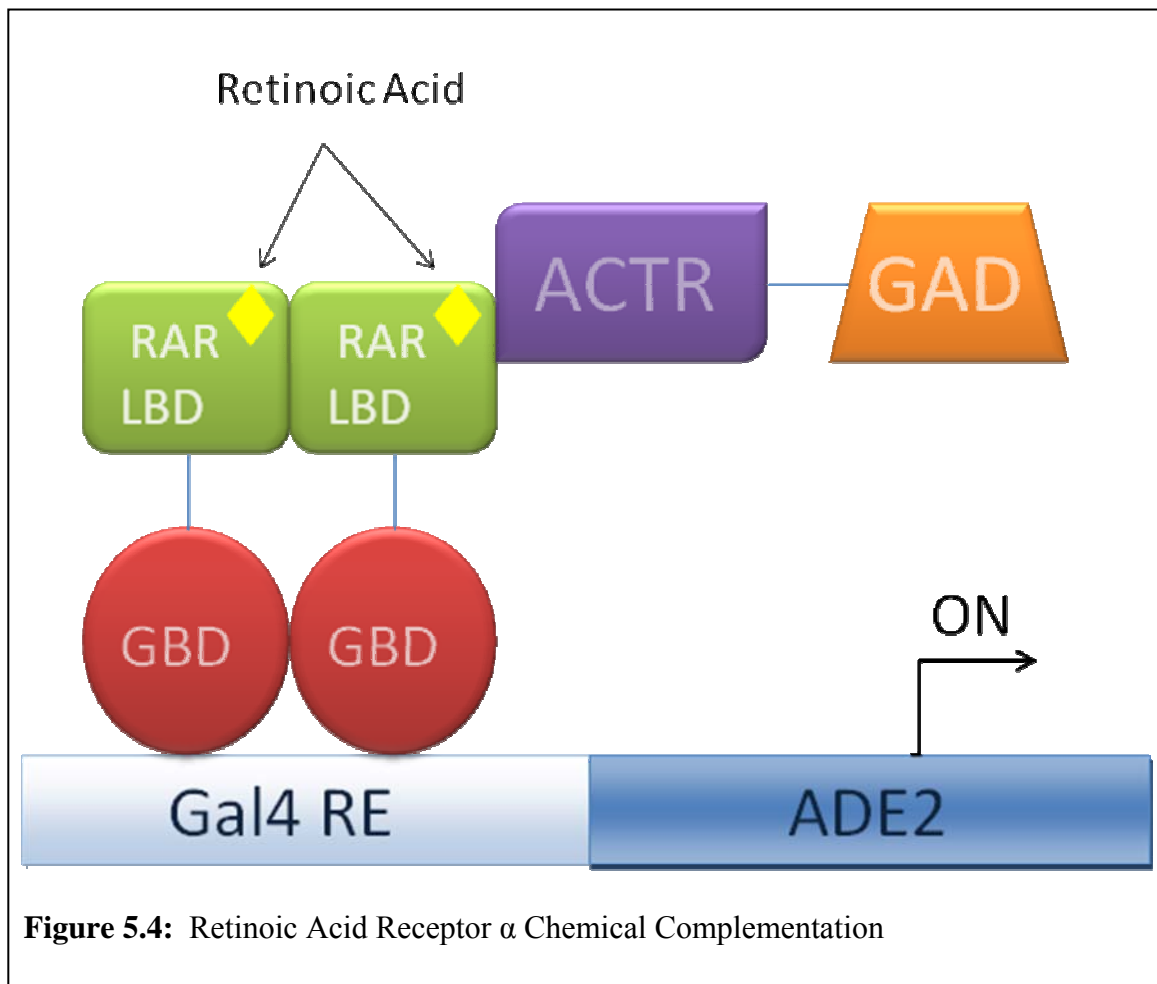
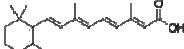
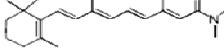
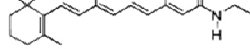


Table 5.2: Charge Reversal Variants Activation by Neutral Amide Ligands

Mutations	sRA	Dimethyl Amide	Ethyl Amide
			
RAR wt	++++	++	++++
R274E	+++	+	+++
K238E	+++	+	+++
S289D	++	—	+
R278E	++	—	+
K229E	+++	+	+++
R274E_K238E	+++	—	++
R274E_S289D	+	—	+
R274E_R278E	+	—	+
R274E_K229E	++	—	++
K238E_S289D	+	—	+
K238E_R278E	+	—	+
K238E_K229E	++	—	++
S289D_R278E	—	—	—
S289D_K229E	+	—	+
R278E_K229E	—	—	—
R274E_K238E_S289D	—	—	—
R274E_K238E_R278E	—	—	—
R274E_K238E_K229E	++	—	+
R274E_S289D_R278E	—	—	—
R274E_S289D_K229E	—	—	—
R274E_R278E_K229E	—	—	—
K238E_S289D_R278E	—	—	—
K238E_S289D_K229E	—	—	—
K238E_R278E_K229E	—	—	—
S289D_R278E_K229E	—	—	—
R274E_K238E_S289D_R278E	—	—	—
R274E_K238E_S289D_K229E	—	—	—
R274E_K238E_R278E_K229E	—	—	—
R274E_S289D_R278E_K229E	—	—	—
K238E_S289D_R278E_K229E	—	—	—
R274E_K238E_S289D_R278E_K229E	—	—	—

++++= activation as low as 0.01 mM

+++= activation as low as 0.1 mM

++= activation as low as 1 mM

+= activation as low as 10 mM

—=dead

Blank= not tested

ligand. In accordance with the hypothesis that the variants are intended to reduce the overall “positive” potential of the pocket believed to stabilize the negatively charged residues, many of our mutants do show a decreased affinity for the negatively charged wild-type ligand in comparison to RARwt. For example, the triple mutant R274E;K236E;K229E has a decreased activity with atRA in comparison to RARwt and atRA. This variant is capable of binding and activating in response to atRA with concentrations as low as 1 μ M, vs. 0.01 μ M with RAR wt (Figure 5.7).

From the results of this library we have observed that RAR α is able to tolerate many of the “charge reversal” mutations. The variants in this library were found to not only activate with neutral RA ligands, but many of these variants, although at a lower sensitivity as compared to RARwt, also retain activity with the natural negatively charged atRA in spite of the disruption of the salt bridges and/or hydrogen bonding between residues, K236E, R278E and S289D, and the carboxylate of atRA. Conversely, the double, triple, quadruple, and quintuple variants that are no longer activated by the wild type ligand were excellent candidates for creating charge reversal ligand-receptor pairs with the positively charge amine based ligands.

5.4.3 Effects of Amine Based Retinoids on Charge Reversal Mutants

In addition to testing the neutral amide ligands, RAR α wild-type (RAR wt) and all variants were tested in CC with the positively charged amine ligands, dimethyl amine (DiAmN RA), ethyl amine (EthAmN RA), and the triethoxysilane amine (SiOEth RA) (Table 5.3). As was seen with atRA and the amide ligands, all of the quadruple and quintuple mutants were inactive with the amine ligands. Additionally, all of the single mutants and many of the double mutants were activated better by atRA than both the

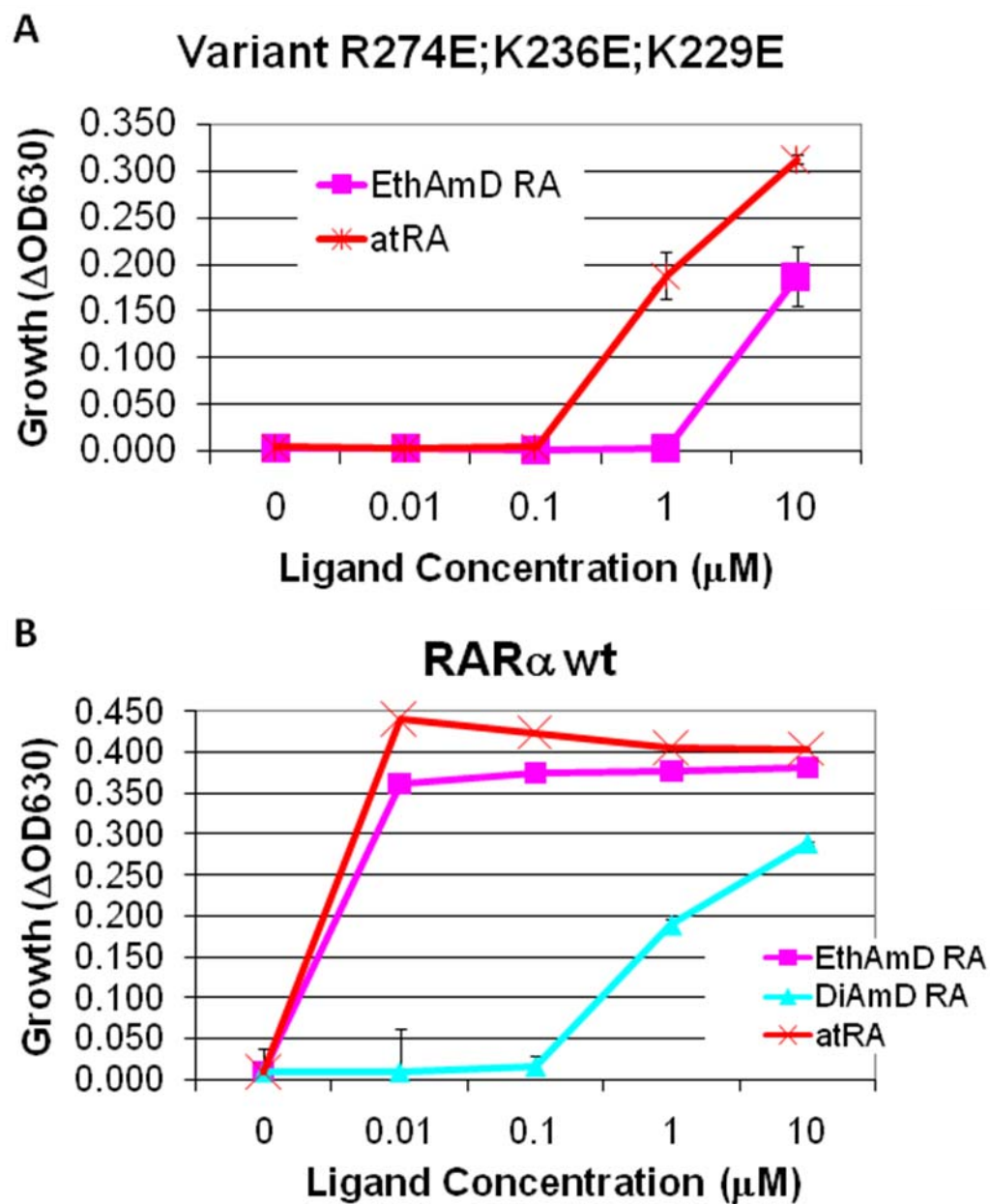


Figure 5.5: (A) Activation of R274E;K236E;K229E Variant with atRA and EthAmD RA
(B) Activation of RARwt with Amide Ligands

DiAmN RA and the EthAmN RA. Interestingly, the single mutant R278E, which is activated by atRA, albeit with a lower affinity than RARwt, shows no activation with the EthAmN RA (Table 5.3). According to Renaud and coworkers, the R278 residues is oriented to form a weak salt bridge with the atRA ligand, but is also a part of the hydrogen bonding network needed to further anchor the carboxylate of atRA (Figure 5.8A) [32]. The inability of the EthAmN RA to activate the R278E variant may be explained by the disruption of this salt bridge and hydrogen bonding network mentioned above. This network is designed to stabilize atRA and when the arginine is modified to a glutamic acid, this network is disrupted decreasing the affinity for this mutant with atRA, which can be seen when we compare RAR wt to the R278E variant's activity with atRA (Figure 5.8B). However, the introduction of the glutamic acid alone does not re-establish a hydrogen bonding network capable of stabilizing the amine in the positively charged ligands. Resultantly, this variant does not stabilize the positive ligand which can explain why no activation is seen with the EthAmN or DiAmN RA. This hypothesis supports Hwang and colleagues argument that wild-type ligand-receptor interactions are stabilized by more than the amino acids in the immediate vicinity of ligand. Altering only local residues leaves these other interactions intact and does not afford additional stability to the reversed salt bridge [30].

Additionally, when comparing the activity of each of the single mutation charge reversal variants, we do observe that variants, R278E and S289D, which altered two of the carboxylate anchoring residues in RARwt, are less active with atRA than any of the other single mutants, confirming that these mutations are more important than the other mutations in stabilizing the ligand (Figure 5.9). Accordingly, the S289D;R278E variant

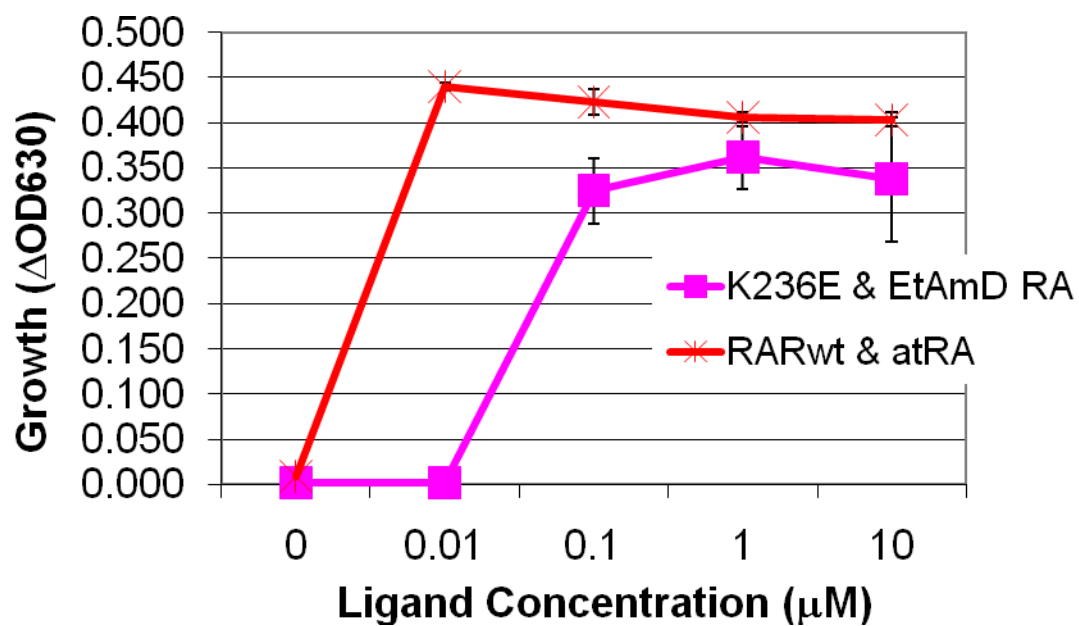


Figure 5.6: Ligand-Receptor Activation Profiles of K236E-EthAmD RA and RARwt-atRA

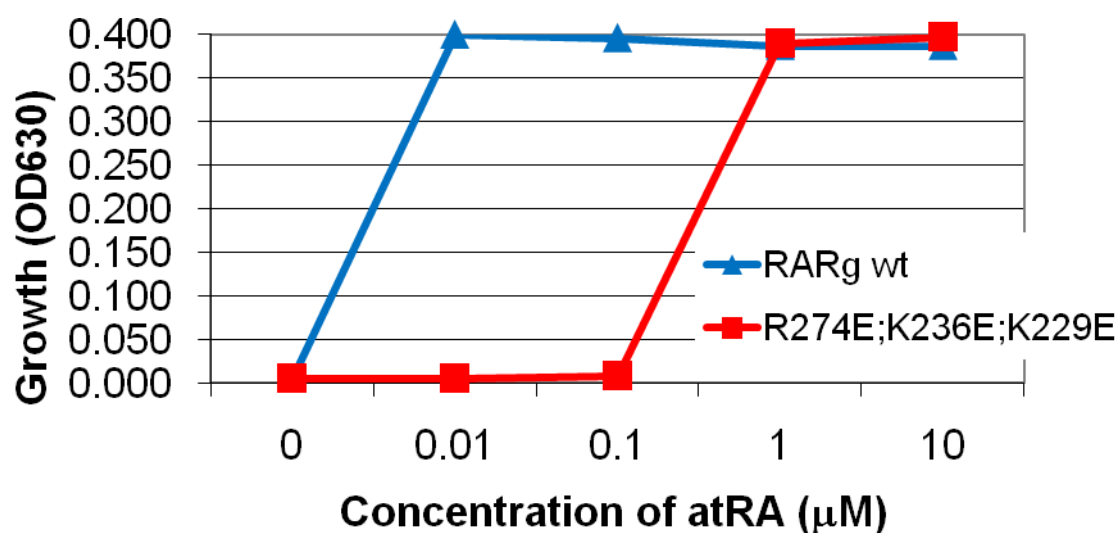
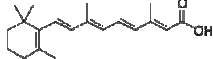
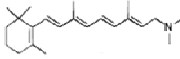
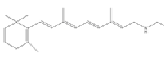



Figure 5.7: RARwt and R274E;K236E;K229E Activation with atRA

Table 5.3: Charge Reversal Variants Activation by Positively Charged Amine Ligands

Mutations	AT-RA 	Dimethyl Amine 	Ethyl Amine 	SIOEtN Amine 
RAR wt	++++	++	+++	++
R274E	+++	++	++	+
K236E	+++	++	++	+
S289D	++	+	+	—
R278E	++	+	—	—
K229E	+++	++	++	+
R274E,K236E	+++	+	+	—
R274E,S289D	+	+	+	—
R274E,R278E	+	+	+	—
R274E,K229E	++	+	+	—
K236E,S289D	+	+	+	—
K236E,R278E	+	—	—	—
K236E,K229E	++	+	—	—
S289D,R278E	—	—	—	—
S289D,K229E	+	—	+	—
R278E,K229E	—	—	—	—
R274E,K236E,S289D	—	—	+	—
R274E,K236E,R278E	—	—	—	—
R274E,K236E,K229E	++	+	+	—
R274E,S289D,R278E	—	—	—	—
R274E,S289D,K229E	—	—	—	—
R274E,R278E,K229E	—	—	—	—
K236E,S289D,R278E	—	—	—	—
K236E,S289D,K229E	—	—	+	—
K236E,R278E,K229E	—	—	—	—
S289D,R278E,K229E	—	—	—	—
R274E,K236E,S289D,R278E	—	—	—	—
R274E,K236E,S289D,K229E	—	—	—	—
R274E,K236E,R278E,K229E	—	—	—	—
R274E,S289D,R278E,K229E	—	—	—	—
K236E,S289D,R278E,K229E	—	—	—	—
R274E,K236E,S289D,R278E,K229E	—	—	—	—

++++= activation as low as 0.01 mM

+++= activation as low as 0.1 mM

++= activation as low as 1 mM

+= activation as low as 10 mM

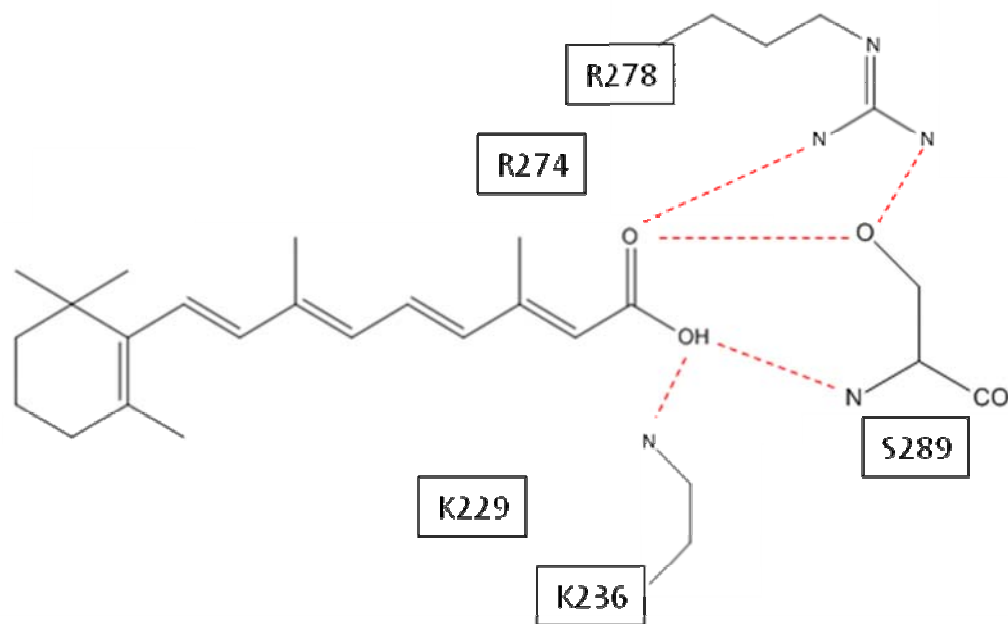
—=dead

Blank= not tested

shows no ligand induced activation with atRA, inferring that this double mutation variant does disrupt the interaction with atRA. However, although these mutations do disrupt the atRA-RAR wt interaction, the newly introduced aspartate and glutamate residues, respectively, do not work to bind and stabilize the positive amine ligands. No activation has been seen with any variant containing the combination of both of these mutations and any ligand. This combination of mutations may render the mutant completely inactive, but may also be a good starting point to introduce a new function into the receptor. There were two interesting triple mutants discovered. The R274E;K236E;S289D variant and the K236E;S289D;K229E variant, which are not activated by either the atRA ligand or the a neutral amide ligands, but are activated by 10 μ M of the positively charged EthAmN RA. Although these variant receptors are orthogonal, binding only the EthAmN RA and not atRA, the EthAmN RA is not orthogonal and activates RARwt as well, and does so better than with the triple mutants. RARwt activates at concentrations as low as 0.1 μ M with the EthAmN RA, whereas the K236E;S289D;K229E and R274E;K236E;S289D variants, only activate at concentrations as low as 10 μ M (Figure 5.10).

When designing the charge reversal library, we hypothesized that we would be able to create an OLRP, in which the charge reversed variants would not only bind the amine ligands in place of the natural atRA ligand, but the positively charged amine ligands would not activate (or would give very low activation) with RARwt, due to the overall charge potential of the binding pocket favoring the negatively charged ligand over the positively charged ligand (Figure 5.11). However, we see a greater activation between the EthAmnN RA and RAR wt than with any of the orthogonal triple mutants

A



B

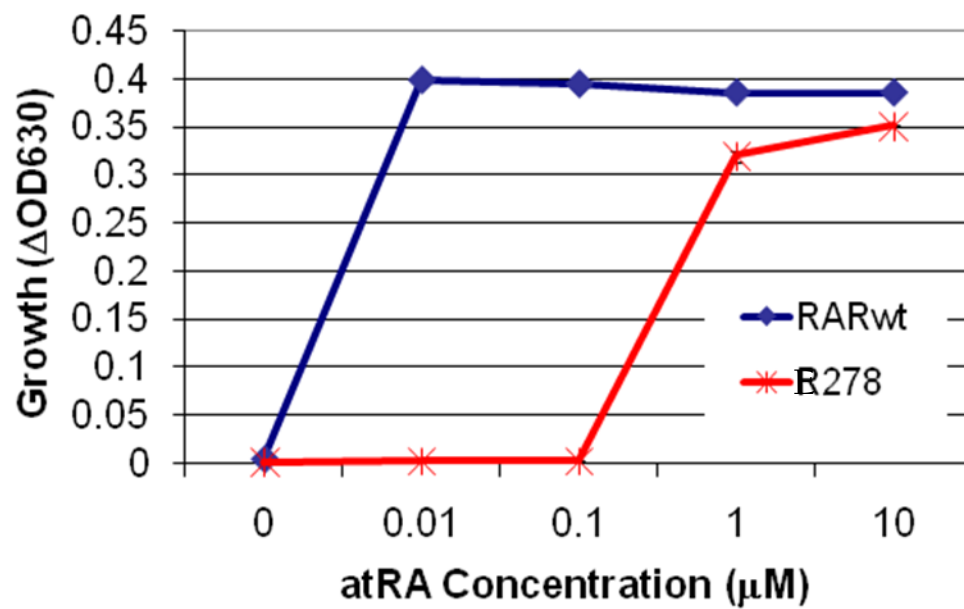


Figure 5.8: (A) Stabilization of atRA by Positively Charged Residues in RAR wt
(B) Activation of RARwt and R278E with atRA

(Figure 5.11). To investigate RARwt's ability to bind the EthAmN RA, we compared docking results of the RARwt-atRA and the RAR wt-EthAmn RA ligand-receptor pairs.

At first inspection of the crystal structure resolved by Renaud and colleagues [32], we noted that the K236 residue had two orientations, the “up” conformation, which pointed towards the solvent, and the “down” conformation, which is pointed into the binding pocket (Figure 5.12). According to Renaud and colleagues, the “down” conformation of K236 forms salt bridges with the carboxylate of atRA, stabilizing the ligand. As a result, we designed our library with the mutation of this residue based on the “down” conformation of the K236. When docking atRA into both the “up” and “down” conformations of RARwt we found that both conformations dock atRA in approximately the same orientation (Figure 5.13A). When in the “down” conformation, K236 forms a salt bridge to stabilize the carboxylate of atRA (Figure 5.13B). However, in the “up” conformation K236 does not have any direct contacts with atRA (Figure 5.13C).

When docking the EthAmN RA ligand into RARwt in both the “down” and “up” conformations of K236, we find that the EthAmN RA docks in almost the same orientation with both K236 conformations (Figure 5.14A). However in the “down” conformation, the amine group is oriented down and away from the K236 residue, as not to clash with this residue, and there are no polar contacts stabilizing the ligand suggesting that it may be unlikely that the ligand would be stable and activate in this conformation (Figure 5.14B). In the “up” conformation of K236, the EthAmN RA orients fairly similarly to atRA. Also in the “up” conformation, although R278 and S289 no longer interact with the ligand to stabilize it, there is a hydrogen bond between L233 and the EthAmN RA, which helps to stabilize the ligand (Figure 5.14C). These results suggest

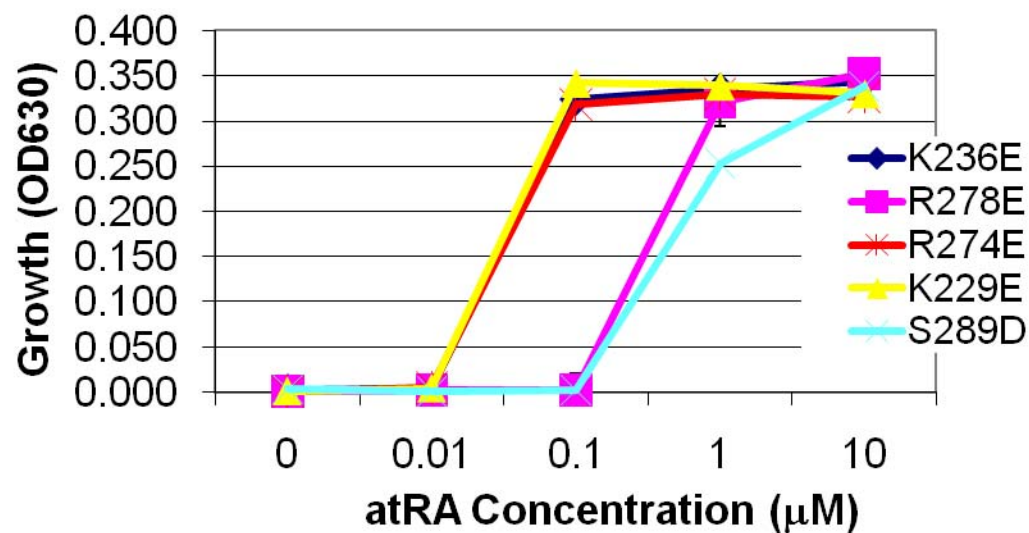


Figure 5.9: Activation of Single Variants by atRA

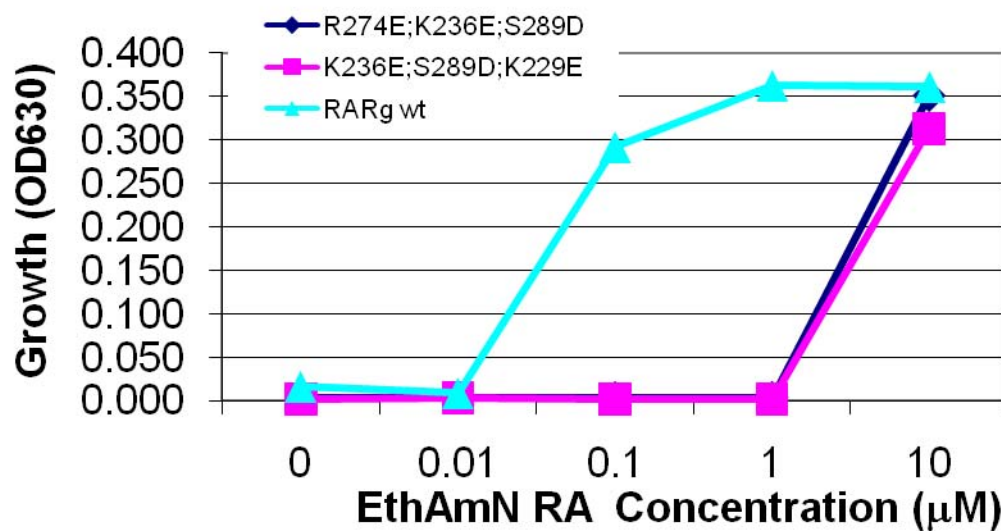


Figure 5.10: Activation of RARwt, R274E;K236E;S289D, and K236E;S289D;K229E Variants by EthAmN RA

that although Renaud and colleagues claim that the K236 residue is oriented in the down conformation helping to stabilize the carboxylate of atRA [32], there is a possibility that the residue is pointing “up” and out of the binding pocket, not making direct contact with the ligand but possibly participating in stabilizing ligand via the cluster of positively charged residues surrounding the pocket. To further validate this claim, all crystal structures of RAR γ , solved with 9cRA and other agonists, orient K236 in the “up” conformation [54-56].

If the K236 residue was oriented in the “down” conformation as Renaud and colleagues first suggested, docking results suggest that the EthAmN RA would not be able to activate RARwt. However, docking results, as well as other crystallographic structures of RAR γ , show that the “up” conformation of K236 is more feasible and would allow the EthAmN to not only bind in place of atRA, but this conformation would allow the ligand to make polar contacts with other residues in the binding pocket to increase stabilization. The experimental results from evaluating the single variants also support the argument that K236 does not directly interact to stabilize atRA. Single variants, R278E and S289D, which are known to directly interact to stabilize atRA, both show a lower sensitivity for atRA as compared to the single variants, R274E and K229E, which do not directly interact with atRA. The K236E variant behaves similar to R274E and K229E single variants and not the R278E and S289D, suggesting that this residue is not directly involved in stabilizing atRA. Ultimately we found that while we were able to create orthogonal triple mutants with the EthAmN RA, the ligand itself was not orthogonal due to its ability to still form polar contacts with other residues in the ligand binding pocket to help stabilize it.

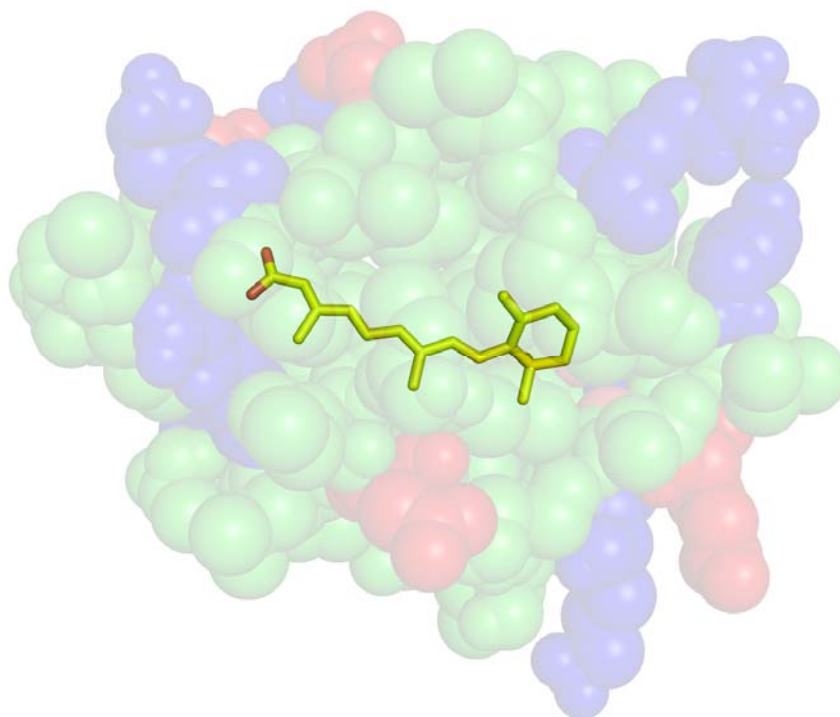


Figure 5.11: Positively Charged Potential Surrounding the Carboxylate of atRA. Neutral residues are shown in green, positively charged residues are shown in blue, and negatively charged residues are shown in red.

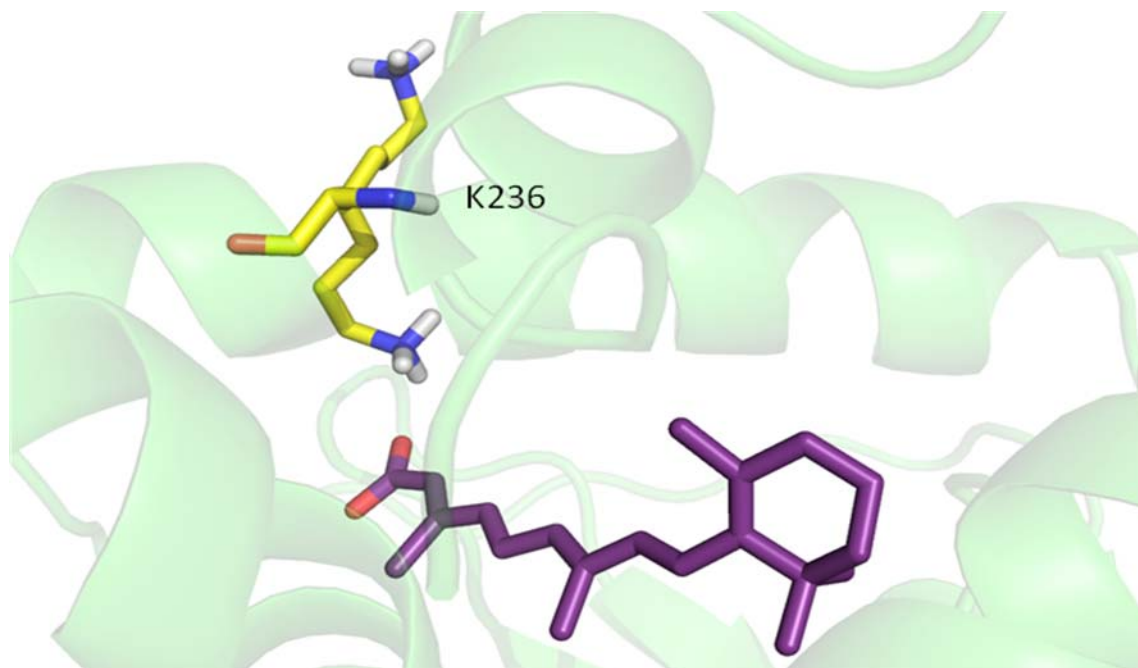


Figure 5.12: “Up” and “Down” Conformations of K236 in RAR γ

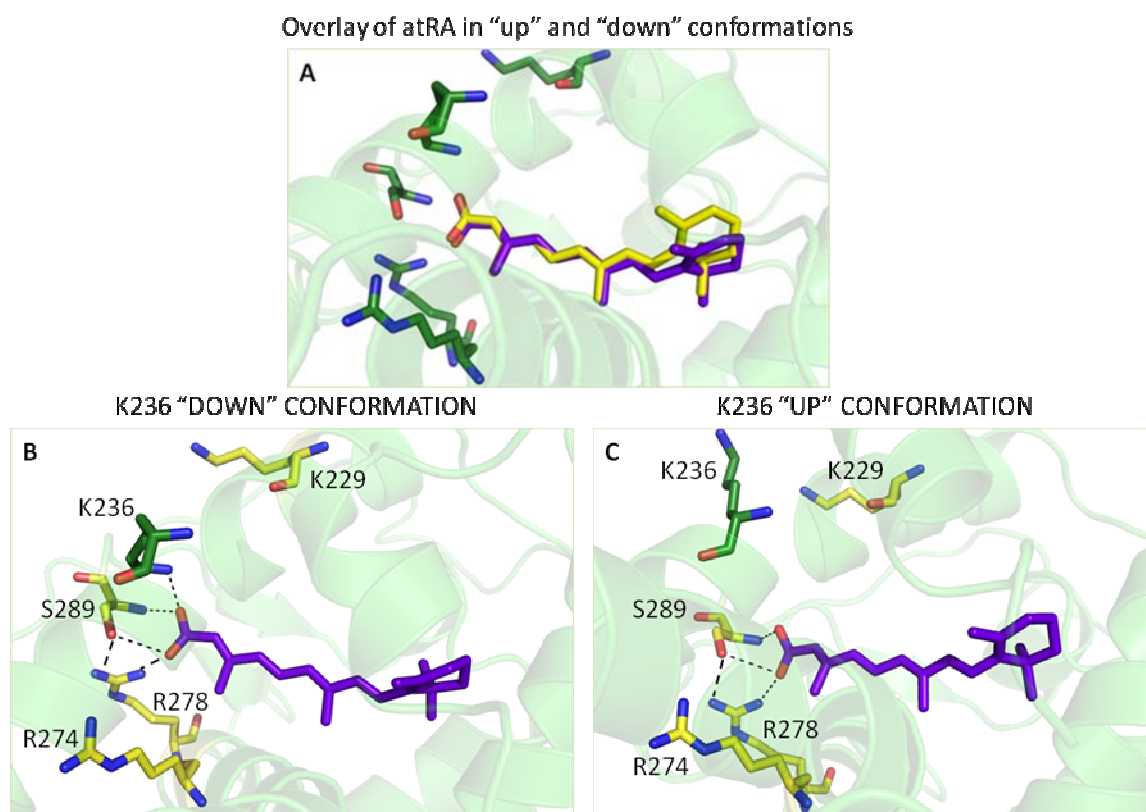


Figure 5.13: Binding of atRA into Binding Pocket of RARwt in “Up” and “Down” Conformations of K236

5.5 Summary and Future Work of Engineering Ligand-Receptor Charge Reversal

Although our library of charge reversal RAR α variants did not include an OLRP or a ligand-receptor pair capable of producing activation potencies equal to or better than RAR α wt and atRA, the library did allow us to gain insight into the structure and function of RAR α . We discovered that the receptor was able to tolerate all the single mutations except the R278E mutation. In addition, we found that R278E and S289D were capable of decreasing the affinity of the receptor for the wild-type atRA ligand better than K236E, K229E, or R274E. Lastly we noted that two triple variants (R274E;K236E;S289D and K236E;S289D;K229E) are capable of shifting the function of the receptor to bind the positively charged EthAmN ligand in place of atRA. Most importantly, we discovered that our results are consistent with Warshel's hypothesis that charge reversal is unlikely to be successful due to energetic barriers.

We are in the process of further characterizing many of the charge reversal variants using docking studies, binding assays, and protein. With this information and the experimental data that we have collected, we will have gained insight into both the structure and function of RAR α and can create variants that better stabilize the positively charged ligands, thus possibly creating variants which are capable of better activating with positively charged amine ligands. Additionally, now that we have a better understanding of RARwt's binding pocket we can design positively charged ligands less likely to activated the wild-type receptor.

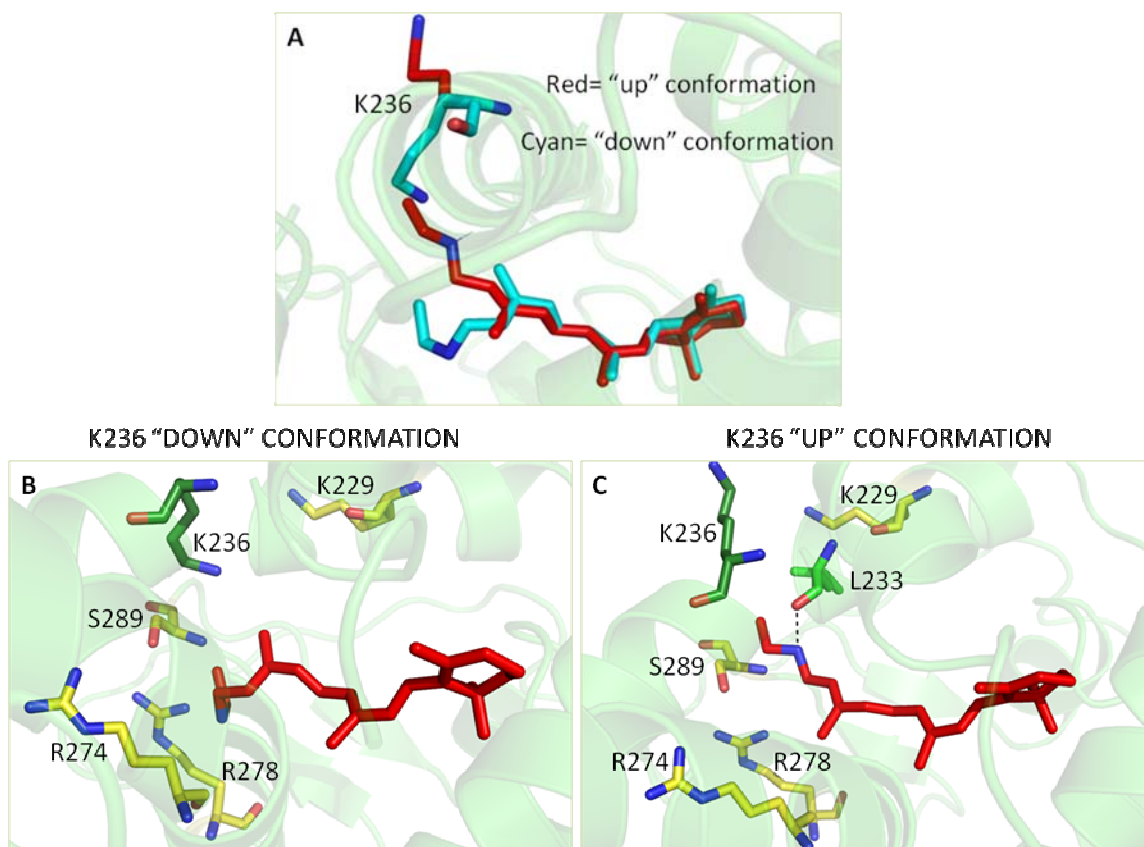


Figure 5.14: Binding of EthAmN RA into Binding Pocket of RARwt in "Up" and "Down" Conformations of K236

5.6 Materials and Methods

Yeast Strain

Yeast strain, PJ69-4A (*MATa trp1-901 leu2-3, 112 ura3-52 his3-200 gal4Δ gal80Δ LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ*), was a kind gift from Dr. Philip James and Dr. Elizabeth Craig (University of Wisconsin, Madison) [57]. The strain contains the *HIS3*, *ADE2*, and *lacZ* genes under the control of different Gal4 promoters.

Ligands

All *trans* retinoic acid (MW=300.44 g/mol) was purchased from Biomol Inc (Plymouth Meeting, PA). The ethyl amide, dimethyl amide, ethyl amine, dimethyl amine, and triethoxysilane retinoids were synthesized by Dr. Stefan France (Georgia Institute of Technology, Georgia). 10 mM stocks of the ligand were dissolved in 80% ethanol:20% DMSO (4:1 v/v) and stored at 4°C.

Expression Plasmids and Site-Directed Mutagenesis

The pGAD10BA ACTR vectors, containing a tryptophan marker, were previously constructed in our lab and contain a Gal4AD fused to the full coactivator, ACTR [53, 58]. To make pGBDRAR α , containing a leucine marker, primers containing a BsmI site at the 5' end and an AflIII site at the 3' end were designed to amplify the LBD of RAR α (containing residues 180-417 of RAR α) from the pCMXRAR α vector, a kind gift from Dr. Ron Evans (Salk Institute for Biological Studies, CA). Once amplified, the gene was ligated into the pGBDRXR vector using the BsmI and AflIII restriction enzymes. Charge reversal mutants were introduced by site-directed mutagenesis using the QuickChange Kit (Stratagene, CA).

Yeast Transformations and Selective Media and Plates

Synthetic complete (SC) plates were made as described previously [59]. Selective media and plates were made of SC media minus, either, leucine and tryptophan or minus adenine, leucine, and tryptophan. pGBDRAR α and pGAD10BA ACTR were introduced into PJ69-4A using the LiAc transformation method [60]. Cells were plated onto SC – Leu-Trp plates and inoculated.

Yeast Chemical Complementation Growth Assays

Yeast quantitation assays were performed as previously described [58]. Briefly, cells were grown in SC-Ade-Leu-Trp media with the appropriate ligands in of 96-well microtiter plate at 30 °C and 150 RPM. OD₆₃₀ was measured at t=0, t=24 hours, and t=48 hours to determine growth in Media lacking adenine.

Docking

AutoDock 4.0 [61, 62] was used to dock atRA and EthAmN RA into the binding pocket of RAR γ . atRA was removed from the PDB structure and redocked into RAR (PDB code 2LBD) [32]. Structure was prepared for Docking using UCSF CHIMERA [63] by removing the ligand and water molecules, adding polar hydrogens, and assigning Kollman united atom charges. Ligands were created using ChemBioDraw Ultra 11.0 and ChemBio3D Ultra 11.0. Then ligands were modified with the AutoDockTools [62] by adding Gasteiger charges. AutoDock4.0 was used to perform docking simulations using default parameters. The K236 residues was altered using UCSF CHIMERA [63]. The structures with the lowest free energy of binding were analyzed.

5.7 Literature Cited

1. Chockalingam K & Zhao H (2005) Creating new specific ligand-receptor pairs for transgene regulation. *Trends Biotechnol* **23**, 333-335.

2. Chockalingam K, Chen Z, Katzenellenbogen JA & Zhao H (2005) Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches. *Proc Natl Acad Sci U S A* **102**, 5691-5696.
3. Panguluri SK, Kumar P & Palli SR (2006) Functional characterization of ecdysone receptor gene switches in mammalian cells. *FEBS J* **273**, 5550-5563.
4. Doyle DF, Braasch DA, Jackson LK, Weiss HE, Boehm MF, Mangelsdorf DJ & Corey DR (2001) Engineering orthogonal ligand-receptor pairs from "near drugs". *J Am Chem Soc* **123**, 11367-11371.
5. Peet DJ, Doyle DF, Corey DR & Mangelsdorf DJ (1998) Engineering novel specificities for ligand-activated transcription in the nuclear hormone receptor RXR. *Chem Biol* **5**, 13-21.
6. Schwimmer LJ, Rohatgi P, Azizi B, Seley KL & Doyle DF (2004) Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules. *Proc Natl Acad Sci U S A* **101**, 14707-14712.
7. Tedesco R, Thomas JA, Katzenellenbogen BS & Katzenellenbogen JA (2001) The estrogen receptor: a structure-based approach to the design of new specific hormone-receptor combinations. *Chem Biol* **8**, 277-287.
8. Hassan AQ & Koh JT (2006) A functionally orthogonal ligand-receptor pair created by targeting the allosteric mechanism of the thyroid hormone receptor. *J Am Chem Soc* **128**, 8868-8874.
9. Koh JT (2002) Engineering selectivity and discrimination into ligand-receptor interfaces. *Chem Biol* **9**, 17-23.
10. Koh JT PM, Tomic-Canic M, McDaniel C (1999) Selective Regulation of Gene Expression Using Rationally-Modified Retinoic Acid Receptors. *Journal of the American Chemical Society* **121**, 1984-1985.
11. Shi Y & Koh JT (2002) Functionally orthogonal ligand-receptor pairs for the selective regulation of gene expression generated by manipulation of charged residues at the ligand-receptor interface of ER alpha and ER beta. *J Am Chem Soc* **124**, 6921-6928.
12. Chambon P (2005) The nuclear receptor superfamily: a personal retrospect on the first two decades. *Mol Endocrinol* **19**, 1418-1428.
13. Gronemeyer H, Gustafsson JA & Laudet V (2004) Principles for modulation of the nuclear receptor superfamily. *Nat Rev Drug Discov* **3**, 950-964.
14. Katzenellenbogen JA & Katzenellenbogen BS (1996) Nuclear hormone receptors: ligand-activated regulators of transcription and diverse cell responses. *Chem Biol* **3**, 529-536.

15. Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schutz G, Umesono K, Blumberg B, Kastner P, Mark M, Chambon P & Evans RM (1995) The nuclear receptor superfamily: the second decade. *Cell* **83**, 835-839.
16. Bishop A, Buzko O, Heyeck-Dumas S, Jung I, Kraybill B, Liu Y, Shah K, Ulrich S, Witucki L, Yang F, Zhang C & Shokat KM (2000) Unnatural ligands for engineered proteins: new tools for chemical genetics. *Annu Rev Biophys Biomol Struct* **29**, 577-606.
17. Germain P, Kammerer S, Perez E, Peluso-Iltis C, Tortolani D, Zusi FC, Starrett J, Lapointe P, Daris JP, Marinier A, de Lera AR, Rochel N & Gronemeyer H (2004) Rational design of RAR-selective ligands revealed by RARbeta crystal structure. *EMBO Rep* **5**, 877-882.
18. Swann SL, Bergh JJ, Farach-Carson MC & Koh JT (2002) Rational design of vitamin D3 analogues which selectively restore activity to a vitamin D receptor mutant associated with rickets. *Org Lett* **4**, 3863-3866.
19. Fersht AR (1985) *Enzyme Structure and Mechanism*, 2nd Ed. edn. Freeman, San Francisco.
20. Belshaw PJaSSL (1997) Cell-Specific Calcineurin Inhibition by a Modified Cyclosporin. *J Am Chem Soc* **119**, 1805-1806.
21. Schneider A, Heyne J & Hessel M (1991) [Non-tropical pyomyositis with fatal outcome]. *Zentralbl Chir* **116**, 1417-1420.
22. Dill KA (1990) Dominant forces in protein folding. *Biochemistry* **29**, 7133-7155.
23. Clarke AR, Atkinson T & Holbrook JJ (1989) From analysis to synthesis: new ligand binding sites on the lactate dehydrogenase framework. Part I. *Trends Biochem Sci* **14**, 101-105.
24. Cronin C MB, and Kirsch J (1987) Reversal of Substrate Charge Specificity by Site-Directed Mutagenesis of Aspartate Aminotransferase. *Journal of the American Chemical Society* **109**, 2222.
25. Hwang YW & Miller DL (1987) A mutation that alters the nucleotide specificity of elongation factor Tu, a GTP regulatory protein. *J Biol Chem* **262**, 13081-13085.
26. Huang F, Buschman LL, Higgins RA & McGaughey WH (1999) Inheritance of resistance to bacillus thuringiensis toxin (Dipel ES) in the european corn borer. *Science* **284**, 965-967.
27. Cronin CN & Kirsch JF (1988) Role of arginine-292 in the substrate specificity of aspartate aminotransferase as examined by site-directed mutagenesis. *Biochemistry* **27**, 4572-4579.

28. Graf L, Craik CS, Patthy A, Rocznik S, Fletterick RJ & Rutter WJ (1987) Selective alteration of substrate specificity by replacement of aspartic acid-189 with lysine in the binding pocket of trypsin. *Biochemistry* **26**, 2616-2623.
29. Kapoor TM & Mitchison TJ (1999) Allele-specific activators and inhibitors for kinesin. *Proc Natl Acad Sci U S A* **96**, 9106-9111.
30. Hwang JK WA (1988) Why ion pair reversal by protein engineering is unlikely to succeed. *Nature* **334**, 270-272.
31. Wells JA, Powers DB, Bott RR, Graycar TP & Estell DA (1987) Designing substrate specificity by protein engineering of electrostatic interactions. *Proc Natl Acad Sci U S A* **84**, 1219-1223.
32. Renaud JP, Rochel N, Ruff M, Vivat V, Chambon P, Gronemeyer H & Moras D (1995) Crystal structure of the RAR-gamma ligand-binding domain bound to all-trans retinoic acid. *Nature* **378**, 681-689.
33. Sporn MB, Dunlop NM, Newton DL & Smith JM (1976) Prevention of chemical carcinogenesis by vitamin A and its synthetic analogs (retinoids). *Fed Proc* **35**, 1332-1338.
34. Kastner P, Mark M & Chambon P (1995) Nonsteroid nuclear receptors: what are genetic studies telling us about their role in real life? *Cell* **83**, 859-869.
35. Blomhoff R (1994) Transport and metabolism of vitamin A. *Nutr Rev* **52**, S13-23.
36. Glass CK & Rosenfeld MG (2000) The coregulator exchange in transcriptional functions of nuclear receptors. *Genes Dev* **14**, 121-141.
37. McKenna NJ & O'Malley BW (2002) Combinatorial control of gene expression by nuclear receptors and coregulators. *Cell* **108**, 465-474.
38. Nagy L, Kao HY, Chakravarti D, Lin RJ, Hassig CA, Ayer DE, Schreiber SL & Evans RM (1997) Nuclear receptor repression mediated by a complex containing SMRT, mSin3A, and histone deacetylase. *Cell* **89**, 373-380.
39. Perissi V & Rosenfeld MG (2005) Controlling nuclear receptors: the circular logic of cofactor cycles. *Nat Rev Mol Cell Biol* **6**, 542-554.
40. Hu X & Lazar MA (2000) Transcriptional repression by nuclear hormone receptors. *Trends Endocrinol Metab* **11**, 6-10.
41. Aranda A & Pascual A (2001) Nuclear hormone receptors and gene expression. *Physiol Rev* **81**, 1269-1304.

42. Lammi J, Perlmann T & Aarnisalo P (2008) Corepressor interaction differentiates the permissive and non-permissive retinoid X receptor heterodimers. *Arch Biochem Biophys* **472**, 105-114.
43. Westin S, Kurokawa R, Nolte RT, Wisely GB, McInerney EM, Rose DW, Milburn MV, Rosenfeld MG & Glass CK (1998) Interactions controlling the assembly of nuclear-receptor heterodimers and co-activators. *Nature* **395**, 199-202.
44. Germain P, Iyer J, Zechel C & Gronemeyer H (2002) Co-regulator recruitment and the mechanism of retinoic acid receptor synergy. *Nature* **415**, 187-192.
45. Dolle P, Ruberte E, Leroy P, Morriss-Kay G & Chambon P (1990) Retinoic acid receptors and cellular retinoid binding proteins. I. A systematic study of their differential pattern of transcription during mouse organogenesis. *Development* **110**, 1133-1151.
46. Tairis N, Gabriel JL, Gyda M, 3rd, Soprano KJ & Soprano DR (1994) Arg269 and Lys220 of retinoic acid receptor-beta are important for the binding of retinoic acid. *J Biol Chem* **269**, 19516-19522.
47. Tairis N, Gabriel JL, Soprano KJ & Soprano DR (1995) Alteration in the retinoid specificity of retinoic acid receptor-beta by site-directed mutagenesis of Arg269 and Lys220. *J Biol Chem* **270**, 18380-18387.
48. Lamour FP, Lardelli P & Apfel CM (1996) Analysis of the ligand-binding domain of human retinoic acid receptor alpha by site-directed mutagenesis. *Mol Cell Biol* **16**, 5386-5392.
49. Zelent A, Guidez F, Melnick A, Waxman S & Licht JD (2001) Translocations of the RARalpha gene in acute promyelocytic leukemia. *Oncogene* **20**, 7186-7203.
50. Chung SS & Wolgemuth DJ (2004) Role of retinoid signaling in the regulation of spermatogenesis. *Cytogenet Genome Res* **105**, 189-202.
51. Jakel H, Fruchart-Najib J & Fruchart JC (2006) Retinoic acid receptor-related orphan receptor alpha as a therapeutic target in the treatment of dyslipidemia and atherosclerosis. *Drug News Perspect* **19**, 91-97.
52. Azizi B, Chang EI & Doyle DF (2003) Chemical complementation: small-molecule-based genetic selection in yeast. *Biochem Biophys Res Commun* **306**, 774-780.
53. Azizi B, Chang EI, Schwimmer LJ & Doyle DF (2009) *In Progress*.
54. Klaholz BP, Renaud JP, Mitschler A, Zusi C, Chambon P, Gronemeyer H & Moras D (1998) Conformational adaptation of agonists to the human nuclear receptor RAR gamma. *Nat Struct Biol* **5**, 199-202.
55. Klaholz B & Moras D (2002) C-H...O hydrogen bonds in the nuclear receptor RARgamma--a potential tool for drug selectivity. *Structure* **10**, 1197-1204.

56. Klaholz BP, Mitschler A & Moras D (2000) Structural basis for isotype selectivity of the human retinoic acid nuclear receptor. *J Mol Biol* **302**, 155-170.
57. James P, Halladay J & Craig EA (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144**, 1425-1436.
58. Azizi B (2005) Chemical complementation a genetic selection system in yeast for drug discovery, protein engineering, and for deciphering and assembling biosynthetic pathways. In *unpublished thesis (PhD)*. Georgia Institute of Technology, Atlanta.
59. Frederick M. Ausubel RB, Robert E. Kingston, David D. Moore, J. G. Seidman, John A. Smith, Kevin Struhl (1999) *Short Protocols in Molecular Biology*, 4th edn. Wiley John & Sons Inc.
60. Guthrie C (2002) Guide to Yeast Genetics and Molecular Cell Biology, Part B In *Methods in Enzymology*. Academic Press.
61. Morris GM, Goodsell, D. S., Halliday, R.S., Huey, R., Hart, W. E., Belew, R. K. and Olson, A. J. (1999) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Computational Chemistry* **19**, 1639-1662.
62. Sanner MF (1999) Python: a programming language for software integration and development. *J Mol Graph Model* **17**, 57-61.
63. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612.

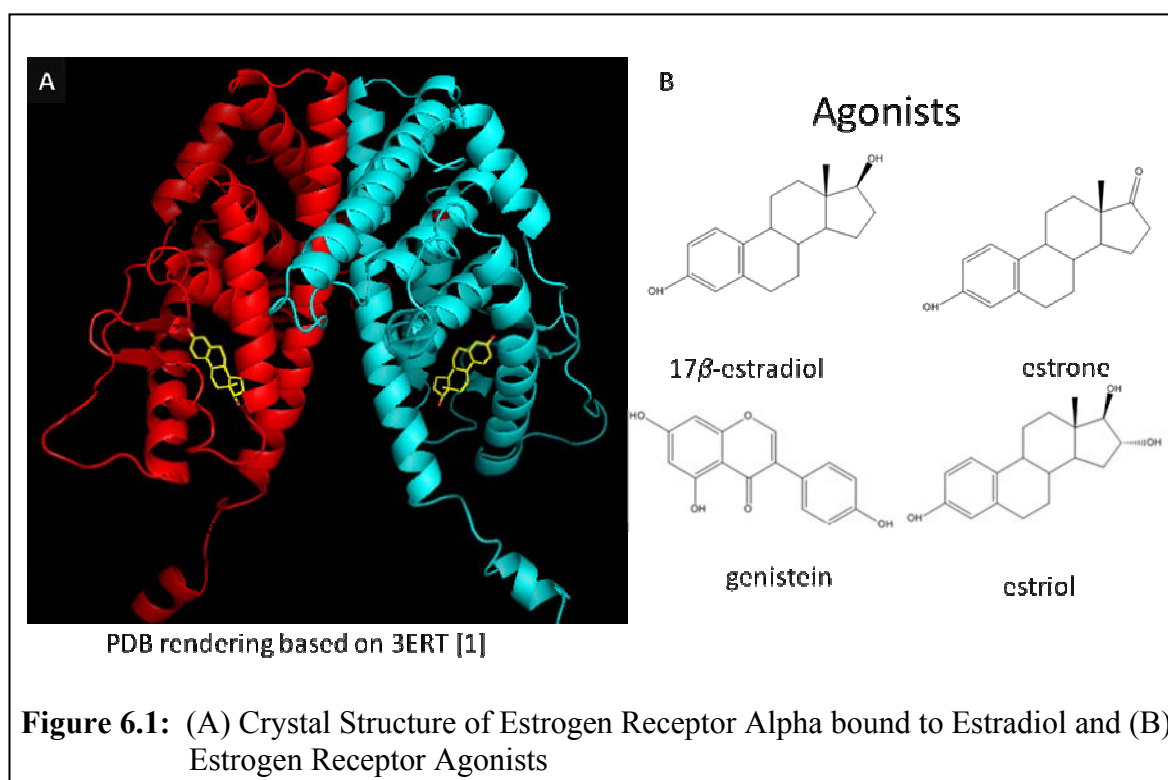
CHAPTER 6

TAMOXIFEN AND HISTONE DEACETYLASE INHIBITOR BASED DUAL INHIBITING COMPOUNDS FOR BREAST CANCER THERAPUTICS

6.1 Estrogen Receptor and Breast Cancer

Breast cancer is the most common cancer in women in industrialized countries [2]. In the United States 1 in 8 women will develop breast cancer at some stage in their lives [3]. The estrogen receptor (ER) is known to play a very crucial role in breast cancer [4]. Over 75% of all breast cancer cases are considered ER positive breast cancers, characterized by over-expression of the ER, specifically ER α [5]. The over-expression of estrogen receptors, which results in an increase in cell proliferation, is proposed to be a result of an increase of cell division and DNA replication in mammary cells in the presence of estrogen. These two processes can cause genetic mutations and contribute to the development of tumors by disrupting DNA repair, the cell cycle, and apoptosis [4, 6]. As a result, drugs that affect the function of ERs are of great interest as breast cancer therapeutics.

Estrogen receptors are nuclear hormone receptors that are activated by the hormone 17 β -estradiol. These receptors are involved in a number of physiological processes, including regulation of proliferation and differentiation in reproductive tract tissues, the cardiovascular system, mammary glands and many other tissues in the body [7-10] (Figure 6.1A). The estrogen receptor is a member of the type I family of nuclear receptors, which are normally located in the cytoplasm and translocated to the nucleus



upon ligand activation [11] (there are also membrane estrogen receptors (mER) that will be discussed briefly later). For nuclear estrogen receptors, this uptake into the nucleus is dependent on a nuclear localization signals (NLS) found in the E domain/LBD of the receptor [12]. Generally this signal is blocked by heat shock proteins (hsp) that bind the receptor until the hormone is present [12]. Once the hormone binds, the receptor undergoes a conformational change in its ligand binding domain. The heat shock proteins dissociate, and the hormone bound receptor enters the nucleus where it homodimerizes, binds DNA, and activates transcription of a target gene.

The estrogen receptor has two isoforms, α and β , that are co-expressed in many of the same cell lines, and as a result can form heterodimers [13]. Although both ERs are widely expressed in different tissue types, there are notable differences between the expression patterns of the two. ER α is found predominantly in breast cancer cells as well as the endometrium, hypothalamus, and ovarian cells. ER β on the other hand is found predominantly in endothelial, brain, heart, kidney, bone, lungs, intestinal, and prostate cells. There are three naturally occurring estrogen ligands in humans known to bind and activate ERs, 17- β estradiol, estrone, and estriol. The natural ligand 17- β estradiol is the most potent of these three estrogen compounds and binds both forms of estrogen, α and β , with the same affinity [14, 15]. However, other ligands have very different affinities for the two receptors. Estrone (also an estrogen) binds preferentially to ER α . ER β binds preferentially to estriol (another estrogen) and genistein (an isoflavone derived from plants) [14, 15] (Figure 6.1B).

In addition to the genomic nuclear estrogen receptors described above, it has also been discovered that there are membrane estrogen receptors (mERs) which aid in

eliciting the regulation of estrogen in a non-genomic, or extranuclear, mechanism [16-20]. It has also been found that the regulation of estrogen by these mERs can cause cell proliferation through rapid signaling of signals such as cAMP and kinase activities [16-18, 21, 22]. In breast cancer cells, this signaling results in crosstalk between the mERs and other proteins, such as the epidermal growth factor receptor (EGFR), which cause activation of the kinase cascades [22-24]. According to O'Malley and colleagues, these kinase cascades can in turn phosphorylate and activate coactivators in the cytoplasm, which then travel to the nucleus and help control nuclear estrogen receptors transcriptional regulation [24].

Selective estrogen receptor modulators (SERMs) are estrogen receptor ligands that have mixed agonist/antagonist functions in various tissue types [25]. These molecules regulate ER's ability to act as a transcriptional activator or repressor in certain cell types by promoting the association of the receptor with either coactivator or corepressor proteins. Therefore, the tissue specificity of these compounds may be associated with the availability of coregulators in specific tissue. The ratio of coactivator to corepressor protein varies in different tissues [26]. A particular ligand may be an agonist in some tissues (where coactivators predominate) and an antagonist in other tissues (where corepressors dominate) [26]. An excellent example of a SERM is the compound tamoxifen. Tamoxifen is generally an ER antagonist in breast cells, and is therefore used in breast cancer treatment [4]. However, this compound is an agonist in bone tissue and the endometrium, aiding in the prevention of osteoporosis but increasing uterine cancer risks [27, 28].

Today, both SERMs and aromatase inhibitors are used to treat breast cancer patients with ER-positive breast cancers [29]. The SERM tamoxifen has been used to treat breast cancer since the 1970's. This compound is the most widely prescribed hormonal agent used to treat breast cancer today [30]. Tamoxifen is a synthetic nonsteroidal antiestrogen that binds to the estrogen receptor and blocking the effects of estrogens, serving as an antagonist (reviewed in Chapter 1). As a result tamoxifen is most effective on estrogen receptor positive breast cancers [30]. Typically, tamoxifen has above a 50 percent curative rate for estrogen receptor positive breast cancers in postmenopausal women, as opposed to approximately 10 percent or less in those with estrogen receptor negative breast cancers [31].

Tamoxifen effectively inhibits ER positive breast cancers by arresting cells in the G0 and G1 phases of the cell cycle. In general, cells in the G0 phase of the cell cycle are easily stimulated to drive the cell cycle into the G1 phase by hormones such as estrogen. This shift from the G0 phase to the G1 phase is crucial in the control of cell proliferation in cancer. To circumvent this effect tamoxifen reverses the effects of estrogen to decrease cell proliferation in breast cancer cells [32]. Tamoxifen prevents pre-cancerous cells from dividing but does not cause cell death, making it cytostatic rather than cytotoxic [33]. Despite its positive effects, tamoxifen has many side effects as a result of its SERM activity [34]. Tamoxifen ideally can work as an antagonist in breast cancer cells while working as an agonist in other estrogen affected cells, having both good and bad effects. For example, while antagonizing breast cancer cells, tamoxifen's ability to act as an agonist in bone cells allows it to prevent bone loss by inhibiting osteoclasts, preventing osteoporosis [35, 36]. Tamoxifen also has the ability to act as a partial agonist

in other tissues, specifically the endometrium [34]. As a result, a significant side effect to tamoxifen is due to the fact that its SERM activity is linked to endometrial as well as uterine cancers in some women.

In addition to the side effects of tamoxifen, patients treated with tamoxifen for ER positive (ER⁺) breast cancer have been known to develop resistance to tamoxifen therapy. Although the majority of breast cancer cases are estrogen receptor positive, there are also estrogen receptor negative (ER⁻) breast cancers, in which proliferation is not associated with over-expression of the estrogen receptor. In these cases hormone therapy is usually not effective resulting in a worse prognosis for patients with ER⁻ breast cancers [37]. As a result, many researchers are developing alternative breast cancer therapeutics for both ER⁺ and ER⁻ breast cancer.

6.2 Motive for Dual Inhibiting Compounds as Breast Cancer Therapeutics

As mentioned previously, tamoxifen is an effective tool to treat breast cancer. Despite its popularity, this small molecule has significant side effects, including contributing to endometrial and uterine cancers [34]. In addition, tamoxifen is also not very effective against estrogen receptor negative breast cancers, and when used therapeutically for up to five years, many previously tamoxifen sensitive breast cancers become tamoxifen resistant, thus requiring alternative therapies. Consequently, new therapeutic strategies, such as aromatase inhibitors and histone deacetylase inhibitors (HDACi) are being developed [38-41]. Aromatase inhibitors block the aromatase enzyme CYP19, which converts androgens to estrogens [42]. These inhibitors essentially “starve” the estrogen receptors from estrogen activation [41]. Aromatase inhibitors are

only used in post-menopausal women, where the major source of estrogen production occurs by means of peripheral tissues [29].

HDAC Inhibitors and Cancer

HDACi are an emerging class of therapeutics for the treatment of cancer because of their ability to arrest the proliferation of nearly all transformed cell types [43-45]. In general, HDACi are compounds that interfere with the function of histone deacetylases (HDACs). HDACs, as well as their counterparts, histone acetyltransferases (HATs), aid in the control of gene expression by altering the interaction between DNA and histones. HATs work to uncoil DNA from histones by adding acetyl groups to their lysines, allowing DNA to be accessible to transcription machinery. HDACs coil DNA around histones by removing the acetyl groups from lysine residues on histone tails, leading to the formation of a condensed and transcriptionally silenced chromatin [46-49]. Although transcriptional control is the predominant function of HDACs, they have also been found to act on nonhistone substrates as well, including transcription factor and coregulator proteins [50]. Altering the acetylation of these proteins would affect both their function and stability, ultimately impacting various cell processes, including cell differentiation, proliferation, and death.

HDACs have been found to remove acetyl groups from histones resulting in a chromatin conformation that prevents the transcription of genes that encode proteins involved in apoptosis [43, 50]. Both natural and synthetic HDACi have been shown to induce cell differentiation, apoptosis, and inhibit angiogenesis in many cancer cells [51-53]. The mechanism by which HDACi promote cell death is complex and has not been fully deciphered. In addition to turning on genes involved in apoptosis, the effects of

HDACi may be a result of the influence of HDAC on nonhistone substrates such as transcription factors. Induced acetylation of these nonhistone proteins can ultimately result in altered transcriptional activity [45].

Various HDACi have been used in clinical trials to treat various cancers, recently the FDA approved the HDACi, suberoylanilide hydroxamic acid (SAHA), to treat T cell lymphoma [54]. Although these inhibitors alone have shown great promise as cancer therapeutics, combination therapy with other anti-cancer drugs seems to be more optimal in obtaining desired therapeutic effects. Preclinical studies have shown that combination therapy has the potential to enhance cell death *in vitro* [55].

HDAC Inhibitors and Breast Cancer

To circumvent the decreased effectiveness of tamoxifen on acquired hormone resistant breast cancers and ER negative breast cancers, much attention has been directed towards HDACi as breast cancer therapeutics with the potential to reverse hormone therapy resistance [56]. The use of HDACi (both alone or in combination with other anticancer agents) to induce cell death in breast cancer has been previously investigated (reviewed in [56]). When tested alone in breast cancer cells, studies have shown that the effects of HDACi were reversible upon drug removal and that clinical concentrations of the HDACi have very little induction of apoptosis [44]. As a result, it has been suggested that HDACi could be more effective when used in combinations with hormone therapy [57-60].

HDACs are known to interact with ERs at various levels, including involvement in the expression of ER itself, in the expression of ER target genes, and in regulation of the binding of heatshock proteins to ERs via their acetylation (Figure 6.2). Researchers

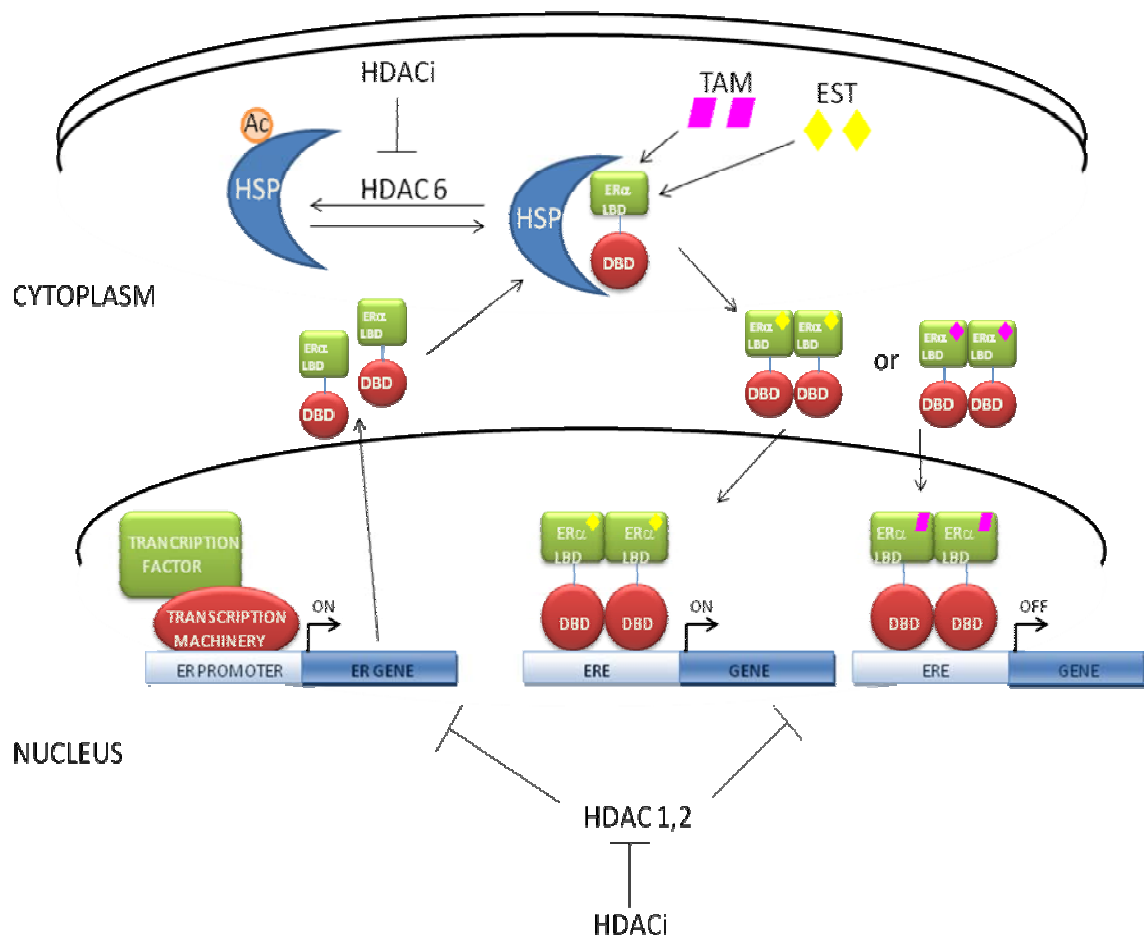


Figure 6.2: HDAC inhibitors, Estrogens, and Anti-estrogen effects on Estrogen signaling

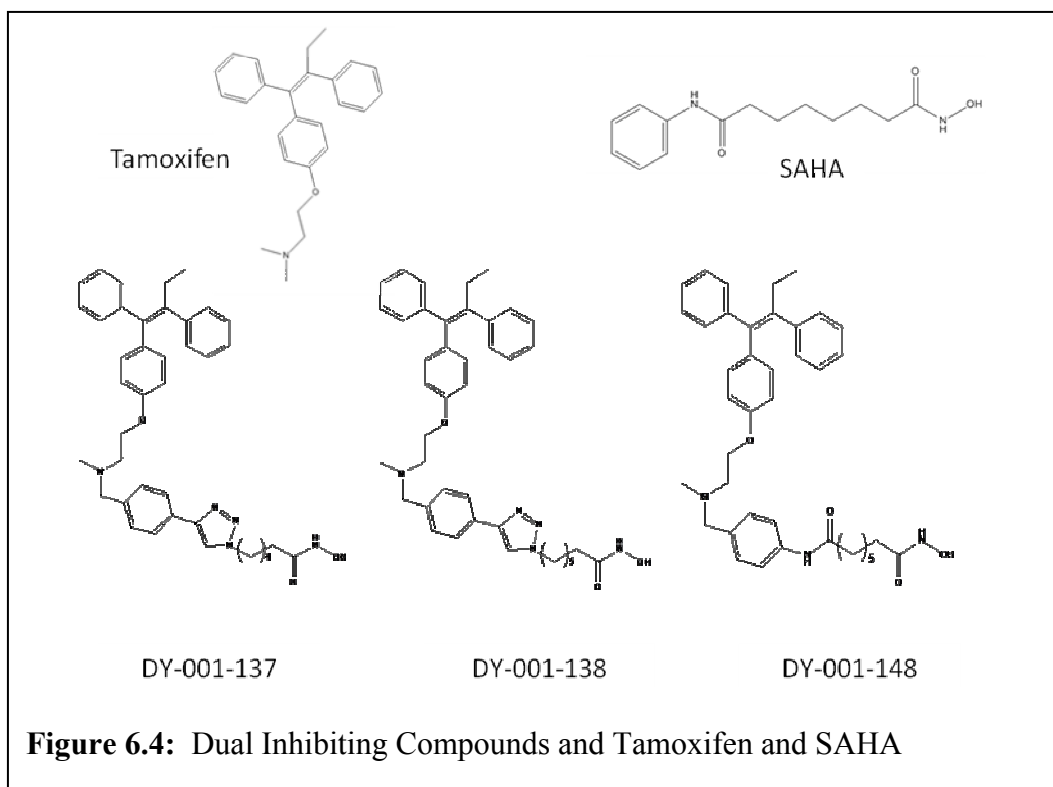
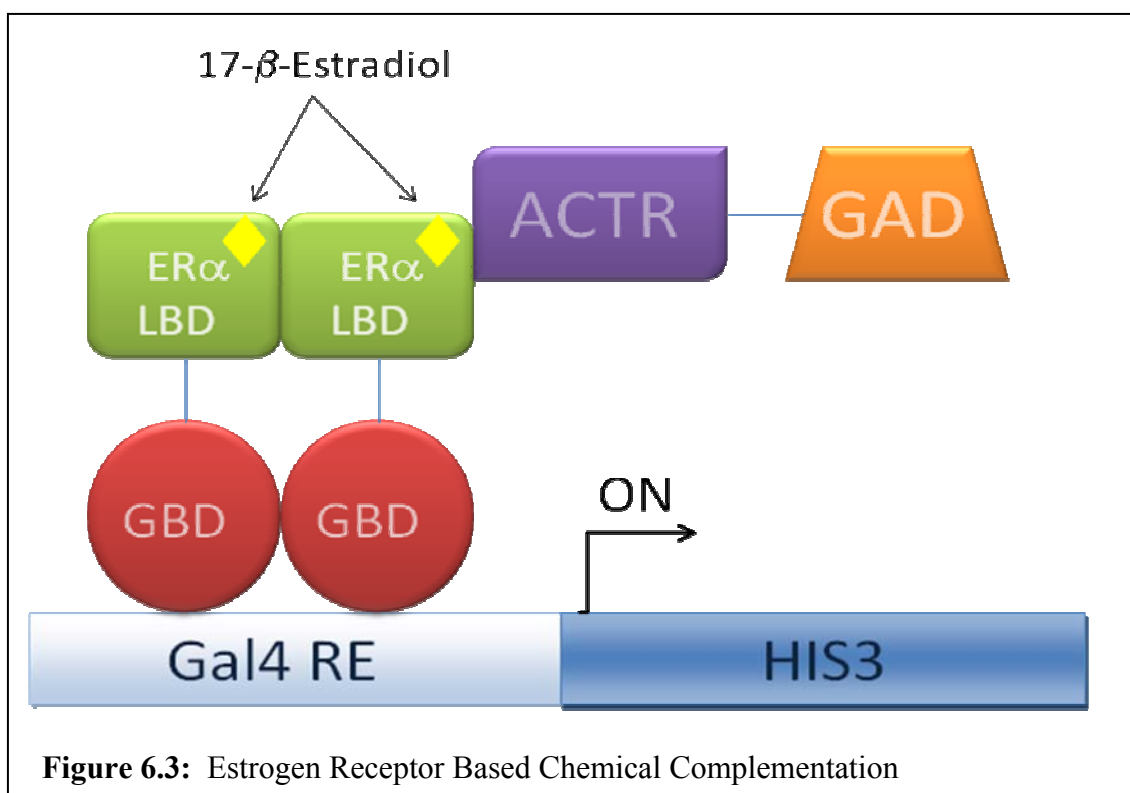
have shown that HDACi sensitize ER negative breast cancer cells and heightens ER positive breast cancer cells to the effects of tamoxifen [57-65]. Kushner and colleagues found that the HDACi, trichostatin A (TSA), SAHA, and valproic acid (VPA) inhibit proliferation of the ER α positive breast cancer cell line (MCF-7), in combination with 10 nM 4-hydroxytamoxifen better than with either agent alone [57]. Jang *et al* reported that in the ER⁻ breast cancer cells, MDA-MB-231, TSA sensitizes the cells to tamoxifen possibly by upregulating ER β expression [58].

To further enhance the effectiveness of HDACi and tamoxifen based combination therapy on breast cancer, it would be beneficial if one could selectively target HDACi into breast cancer cells. Most HDACi, to include SAHA, are considered pan-inhibitors in that they non-selectively inhibit class I, II, and IV HDAC enzymes, which includes 11 of the 18 known human HDACs found throughout various cells and tissues in the body [44, 45]. The ability to inhibit such a broad classes of HDACs can produce side effects, such as cardiac toxicity [66-69], and reduced potencies with these HDACi. To undertake the broad inhibition, we aim to develop a breast cancer therapeutic that will selectively target HDACi into breast cancer cells by covalently linking HDACi moieties to SERMs moieties. By creating these dual inhibiting conjugates, we can potentially sequester HDACi into breast cancer cells via the binding of the SERM moiety of the conjugates to the ER α (since the ER α is present in 75% of breast cancers). Additionally, because the dual inhibiting compounds contain independently anti-proliferative moieties, the SERM tamoxifen and an HDACi, these compounds are expected to act across various stages of the cancer cell cycle, targeting a larger population of cells, resulting in superior anti-

proliferative activity in comparison to existing agents. From these studies, our hypothesis is that a single compound that covalently links HDACi-like moieties to tamoxifen, should enhance the effectiveness and potency of our dual inhibiting compound in comparison to either of these inhibitors administered independently. The Oyelere lab at Georgia Institute of Technology has designed and synthesized dual inhibiting HDACi-SERM conjugates that combine tamoxifen based moieties with the HDACi, SAHA, like moieties, referred to as DY-001-137, DY-001-138, and DY-001-148. The remainder of this chapter will focus upon the effectiveness of these dual compounds on ER α activity and the anti-proliferative effects in ER α positive and negative breast cancer cell lines.

6.3 Using Chemical Complementation to Assess Activity of Dual Inhibiting Compounds towards Estrogen Receptor

To test whether our dual inhibiting compounds affect the transcriptional activity of ER α , the compounds were assayed using chemical complementation (CC). As mentioned in Chapter 2, CC is a generalizable selection system in *S. cerevisiae* used as a tool to investigate protein-small molecule interactions [70, 71]. This system uses genetic selection to link the survival of a yeast cell to the presence of a specific small molecule via estrogen receptor alpha. The CC system used here to test the effects of the dual inhibiting compounds was constructed using the following two fusion proteins: The Gal4 DBD (GBD) was fused to the ER α ligand binding domain, ER α LBD, (GBD: ER α LBD) and the human nuclear receptor co-activator, ACTR, was fused to the yeast co-activator Gal4's activation domain, GAD (ACTR:GAD). The GBD binds the Gal4 response element, regulating the transcription of the *HIS3* gene. Upon binding of an agonist ligand, such as 17- β -estradiol, the LBD of the nuclear receptor undergoes a



conformational change, recruiting the co-activator, ACTR:GAD, fusion protein. This event initiates transcription of the *HIS3* gene allowing cells containing this system to grow in media lacking histidine (Figure 6.3). Transcription of the *HIS3* gene occurs only when an agonist can bind to the nuclear receptor's LBD and recruits the co-activator fusion protein, required to initiate transcription. Without activation of the nuclear receptor with an agonist, no transcription will occur and yeast will not grow in media lacking histidine [70, 72].

The three dual inhibiting compounds, DY-001-137, DY-001-138, DY-001-148 (Figure 6.4), were tested in CC for their ability to induce ligand activated growth. Results show that while estradiol is capable of inducing ligand activated growth in CC, with an half maximal effective concentration (EC_{50}) value of approximately 300 pM, none of the dual inhibiting compounds are capable of producing ligand activated growth at any ligand concentration (Figure 6.5). CC can also be used to assess whether the small molecule is an agonist or an antagonist. Since the compounds were found not to be agonists of ER α , to assess if the dual inhibiting compounds maintained the antagonist activity of tamoxifen, CC was employed using 300 pM 17- β -estradiol, the EC_{50} concentration, and various concentrations of the dual inhibiting compounds. As mention earlier, EC_{50} concentrations of 17- β -estradiol cause transcription of the *HIS3* gene, thus growth of cells in media lacking histidine. However, if compounds are antagonist, they will be capable of binding and displacing the agonist, 17- β -estradiol, not allowing transcription to occur, thus decrease the growth observed as a result of agonist activity (Figure 6.6). In CC, the dual inhibiting compounds were not only capable of displacing

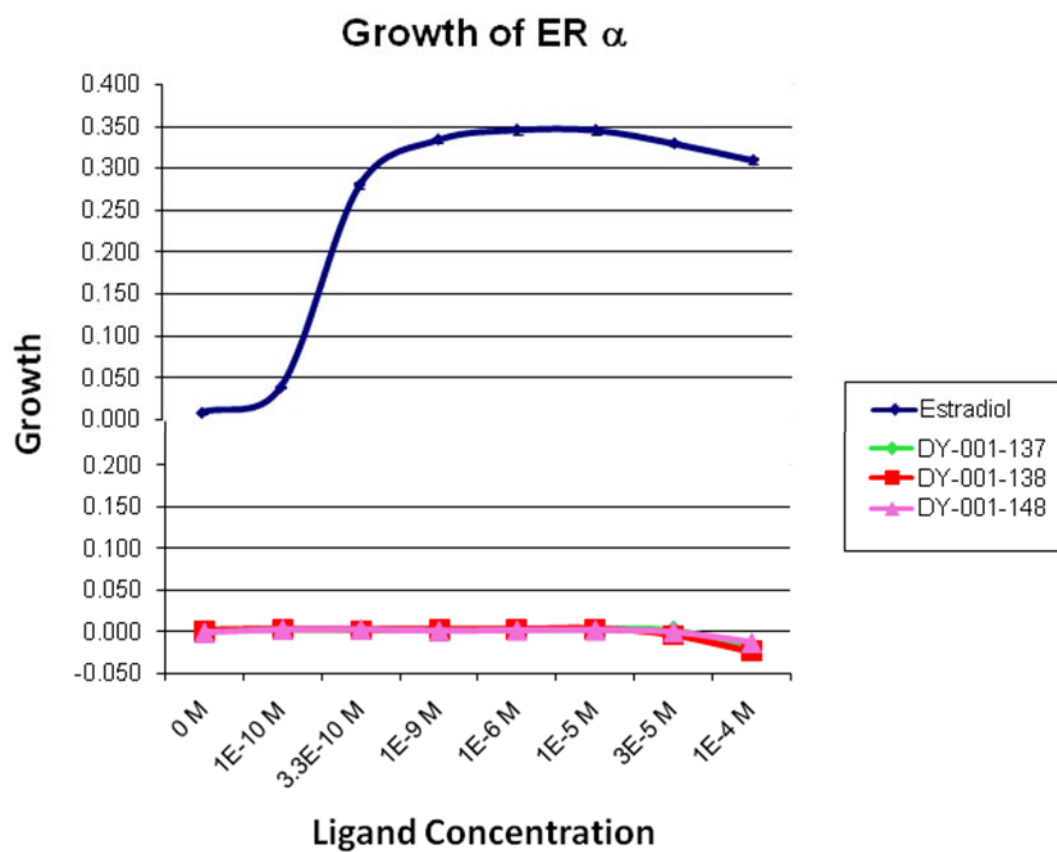


Figure 6.5: Growth Dose Responses of Estrogen Receptor alpha with Estradiol, DY-001-137, DY-001-138, and DY-001-148

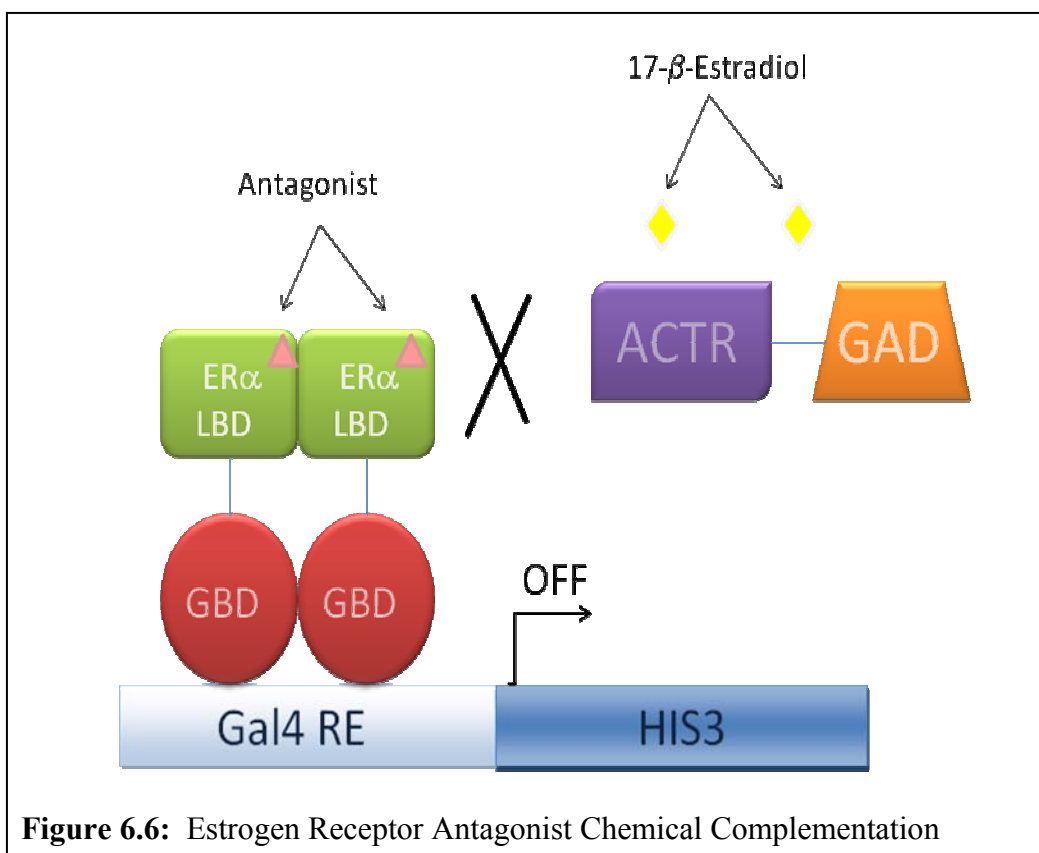


Figure 6.6: Estrogen Receptor Antagonist Chemical Complementation

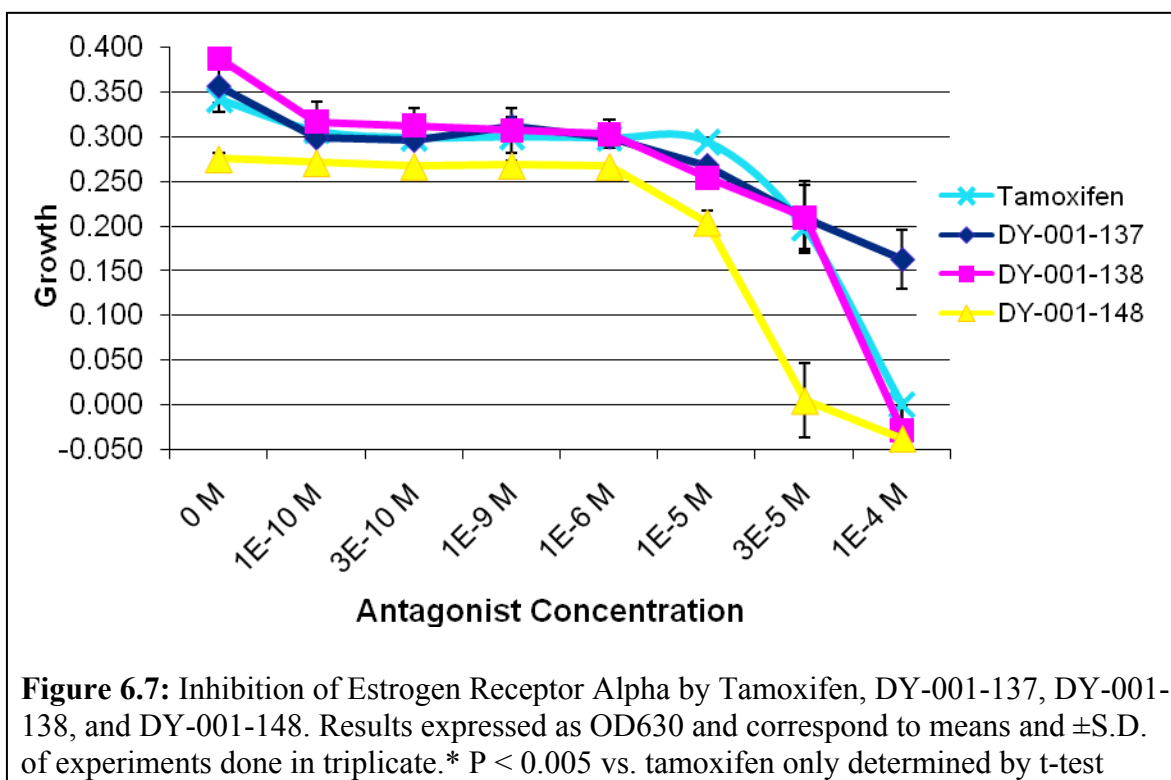


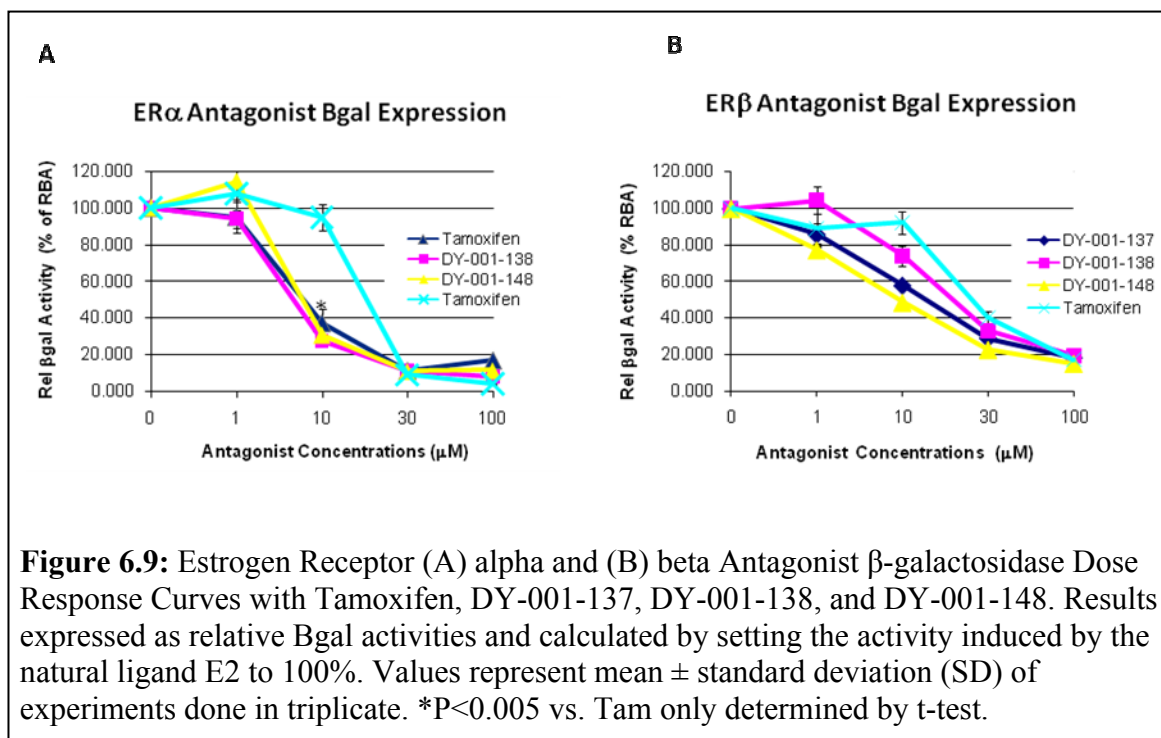
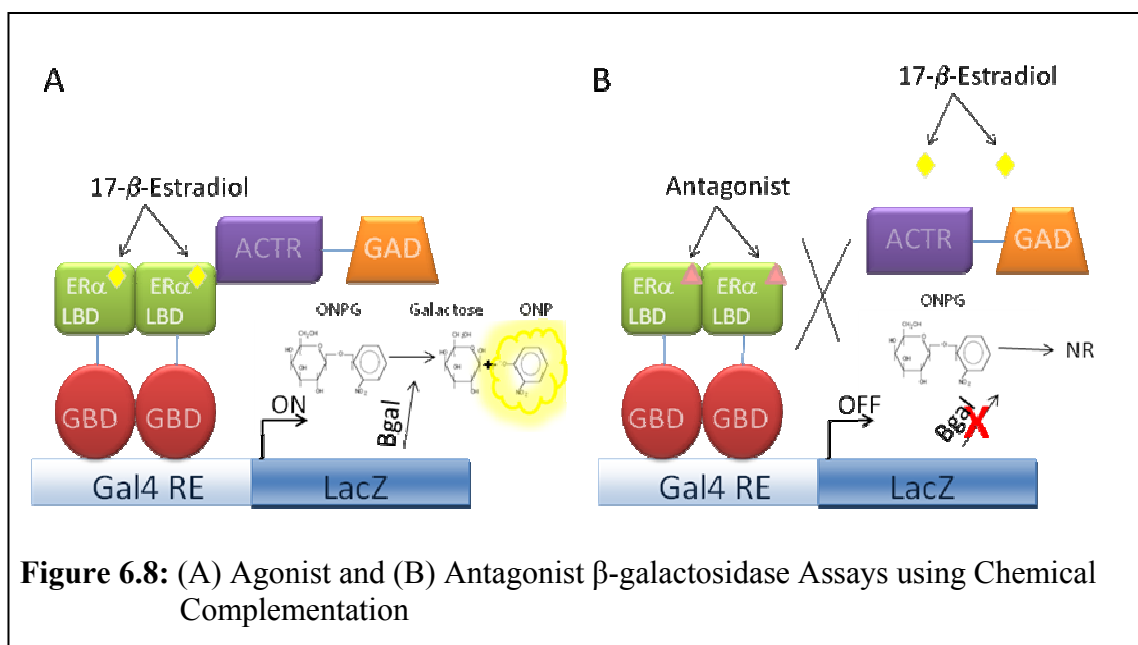
Figure 6.7: Inhibition of Estrogen Receptor Alpha by Tamoxifen, DY-001-137, DY-001-138, and DY-001-148. Results expressed as OD630 and correspond to means and \pm S.D. of experiments done in triplicate. * $P < 0.005$ vs. tamoxifen only determined by t-test

estradiol, but they also behave as a better antagonist than tamoxifen, beginning to inhibit estradiol induced growth with 10 μ M and completely decreasing estradiol induced growth at 30 μ M concentrations as opposed to 100 μ M concentrations necessary for tamoxifen antagonism (Figure 6.7).

To quantify the decrease in ER α transcriptional activity in the presence of these compounds, the *lacZ* gene was utilized in place of the *HIS3* gene in CC. *LacZ* encodes the β -galactosidase (Bgal) enzyme. When o-nitrophenyl β -D-galactopyranoside (ONPG) is added to the media, Bgal catalyzes the conversion of ONPG to galactose, a colorless compound, to o-nitrophenol (ONP), a yellow compound. The amount of ONP produced can be measured by determining the absorbance at 405 nm (Figure 6.8). These results correlate with those of the growth assays. The dual inhibiting compounds display antagonist activity at 10 μ M, resulting in a decrease in expression of the *lacZ* gene in the presence of 300 pM estradiol from 100% of the maximum response, when no dual inhibiting compound is present, to approximately 30% at 10 μ M concentrations of the compounds.. At 10 μ M, tamoxifen does not antagonize the effects of estradiol. Tamoxifen requires 30 μ M concentrations to begin to antagonize the effects of estradiol (Figure 6.9A). Results with ER β are similar to those seen with ER α (Figure 6.9B). The dual inhibiting compounds are also more potent than tamoxifen with ER α in yeast, displaying half maximal inhibitory concentrations (IC₅₀) of 6 μ M as opposed to 18 μ M with tamoxifen.

6.4 Whole Cell Proliferation Assays in Estrogen Receptor positive and negative cells

To determine whether the dual inhibiting compounds are able to decrease



proliferation in breast cancer cells, the compounds were tested in the ER α positive and negative cell lines, MCF-7 and MDA-MB-231, respectively, for anti-proliferative activity. The affinities of the compounds were compared to those of tamoxifen alone, SAHA alone, and equal amounts of tamoxifen and SAHA in both cell lines. To assess the ability of the each compound to decrease the proliferation of breast cancer cells, cells were grown in the presence of respective compounds and cell viability was evaluated using the tetrazolium compound, MTS [3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium)]. When MTS is introduced into the cell media, living cells are able to bioreduce this compound, using NADPH in a reaction catalyzed by a dehydrogenase enzymes, into a soluble colored formazan product that can be quantified by taking the absorbance of the solution at 490 nm (Figure 6.10A).

6.4.1 Effects of Dual Inhibiting Compounds in ER α positive Breast Cancer Cells

The MCF-7 cells line was derived from a patient with metastatic breast cancer and has since been considered a well-characterized ER α positive control cell line useful for the study of the role of hormone therapy in breast cancer. To test the effects of our dual inhibiting compounds in the MCF-7 cell line, cells were cultured in 96 well plates for two days before adding the dual inhibiting compounds, tamoxifen, or SAHA (at concentrations represented in the results). The cells were treated with each drug for 72 hours and effects on proliferation were evaluated using the MTS proliferation assay. When comparing the anti-proliferative effects of the dual inhibiting compounds, DY-001-137, DY-001-138, and DY-001-148, to those of tamoxifen, we observed that all dual inhibiting compounds are more potent than tamoxifen in MCF-7 cells with IC_{50} values of 17 μ M for DY-001-137, 14 μ M for DY-001-138, and 7 μ M for DY-001-148 compared to

Table 6.1: Half Maximal Inhibitory Concentrations (IC₅₀) of Compounds in ER⁺ and ER⁻ Cancer Cell Lines

Compound	MCF-7 (ER ⁺)	MDA-MB-231 (ER ⁻)
	IC ₅₀ (μM) ^a	IC ₅₀ (μM) ^a
TAM	24 ± 1.0	24 ± 0.4
SAHA	17 ± 4.0	13 ± 0.2
DY-001-148 ^{b,c,d}	7 ± 0.3	8 ± 0.3
DY-001-138 ^b	15 ± 3.0	21 ± 1.0
DY-001-137 ^b	17 ± 0.8	22 ± 4.0
TAM/SAHA	14 ± 0.8	10 ± 1.0

Values are means (±SEM) from two to three experiments done in triplicate

^a Inhibitory concentration as obtained by MTS assay

^b P < 0.005 vs. TAM determined by t-test analysis for MCF-7

^c P < 0.05 vs. SAHA determined by t-test analysis for MCF-7

^d P < 0.005 vs. SAHA determined by t-test analysis for MCF-7

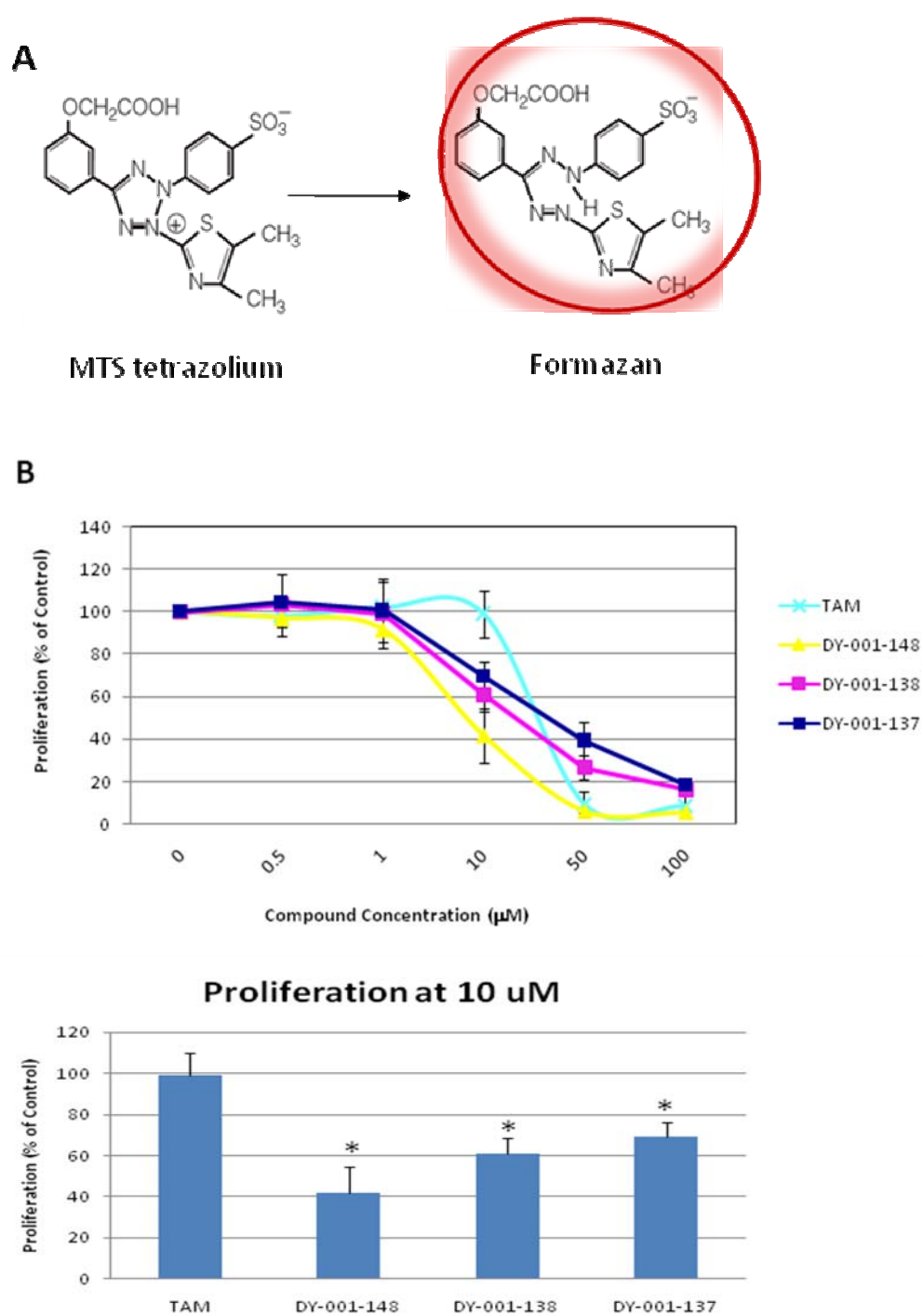


Figure 6.10: Effects of Tamoxifen, DY-001-137, DY-001-138, and DY-001-148 on viability of MCF-7 Cells Using MTS Cell Proliferation Assay. (A) MTS Assay (B) Proliferation Dose Response Curves and Proliferation at 10 μM ligand. Results correspond to means and \pm S.D. of two to three experiments done in triplicate. * $P < 0.005$ vs. Tam alone determined by t-test

24 μ M for tamoxifen ($P \leq .005$) (Table 6.1). In addition to the enhanced potency, we also observed that at 10 μ M, the dual inhibiting compounds have an enhanced efficacy in comparison to tamoxifen. Tamoxifen does not decrease cell proliferation at this concentration, but with DY-001-148 a 60% decrease in cell proliferation was observed, a 40% decrease was observed with DY-001-138, and a 30% decrease with DY-001-137 (Figure 6.10B). Furthermore, the DY-001-148 compound is found to have an enhanced potency in comparison to SAHA with an IC_{50} value of 7 μ M with DY-001-148 vs. 17 μ M with SAHA (Table 6.1). Additionally, at concentrations above 10 μ M, DY-001-148 has better anti-proliferative activity than SAHA, decreasing proliferation by approximately 95% when using 50 to 100 μ M concentrations of DY-001-148, and only by approximately 60% when using that same concentration of SAHA (Figure 6.11).

We also investigated if covalently linking the SAHA-like and tamoxifen-like moieties, as seen in the dual inhibiting compounds, was more effective than if we were to introduce equal amounts of SAHA and tamoxifen into the cell (not covalently linked). We found that at 10 μ M, DY-001-148 is two times more potent than adding non-covalently linked tamoxifen and SAHA and has an enhanced efficacy in comparison to non-covalently linked tamoxifen and SAHA, resulting in a IC_{50} of 7 μ M and a 60% decrease in cell proliferation with the covalently linked DY-001-148 compound vs. a IC_{50} of 14 μ M and a 40% with noncovalently linked tamoxifen and SAHA ($P \leq .005$) (Table 6.1 and Figure 6.12). DY-001-137 and DY-001-138 show a lower efficacy than when adding equal amounts of tamoxifen and SAHA. This may be a result of the variations in the SAHA like moieties of these compounds as compared to SAHA itself (Figure 6.4). Whereas, DY-001-137 and DY-001-138 have slight variations in their SAHA-like

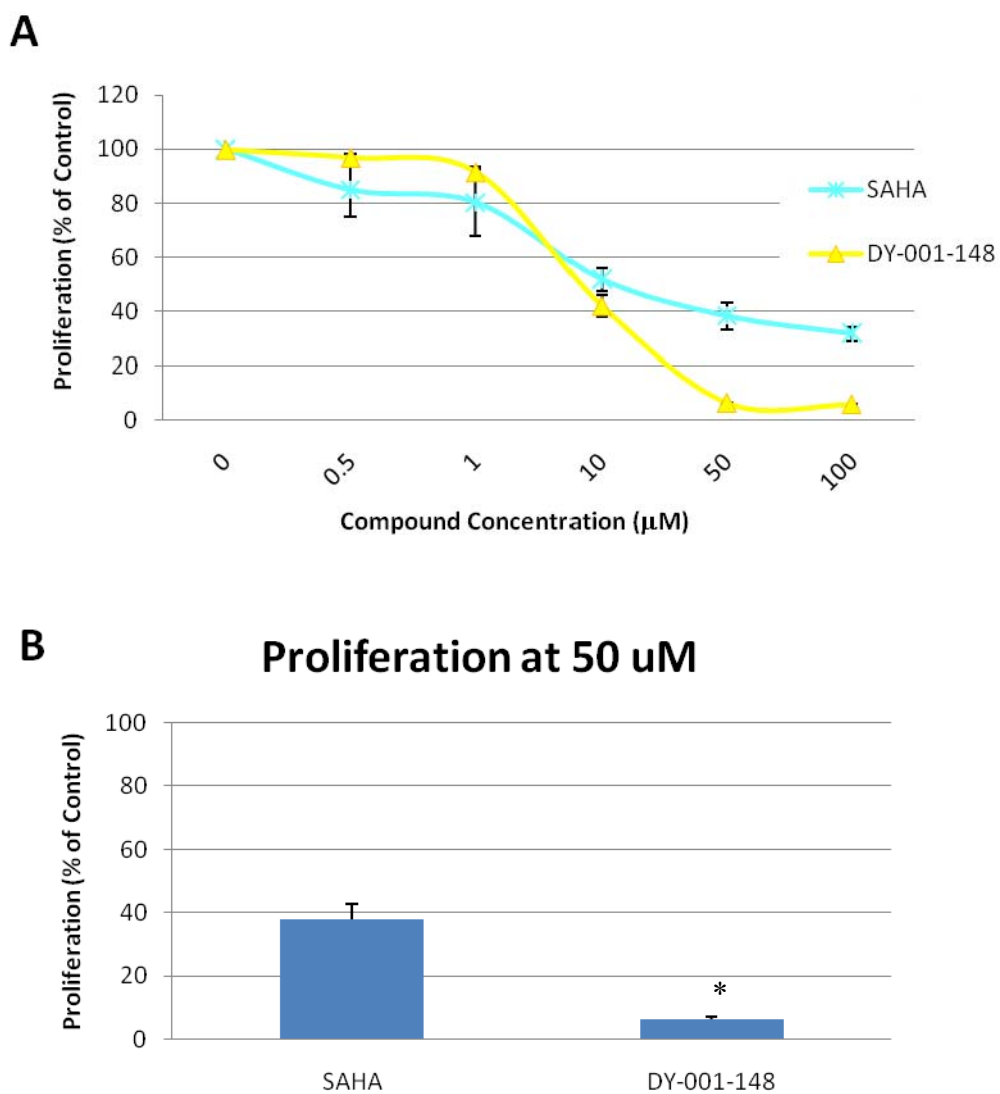


Figure 6.11: Proliferation of MCF-7 Cells in the presence of SAHA and DY-001-148. (A) Dose Response Curves and (B) Proliferation at 50 μM ligand. Results correspond to means and \pm S.E. of two to three experiments done in triplicate. * $P < 0.005$ vs. SAHA alone determined by t-test

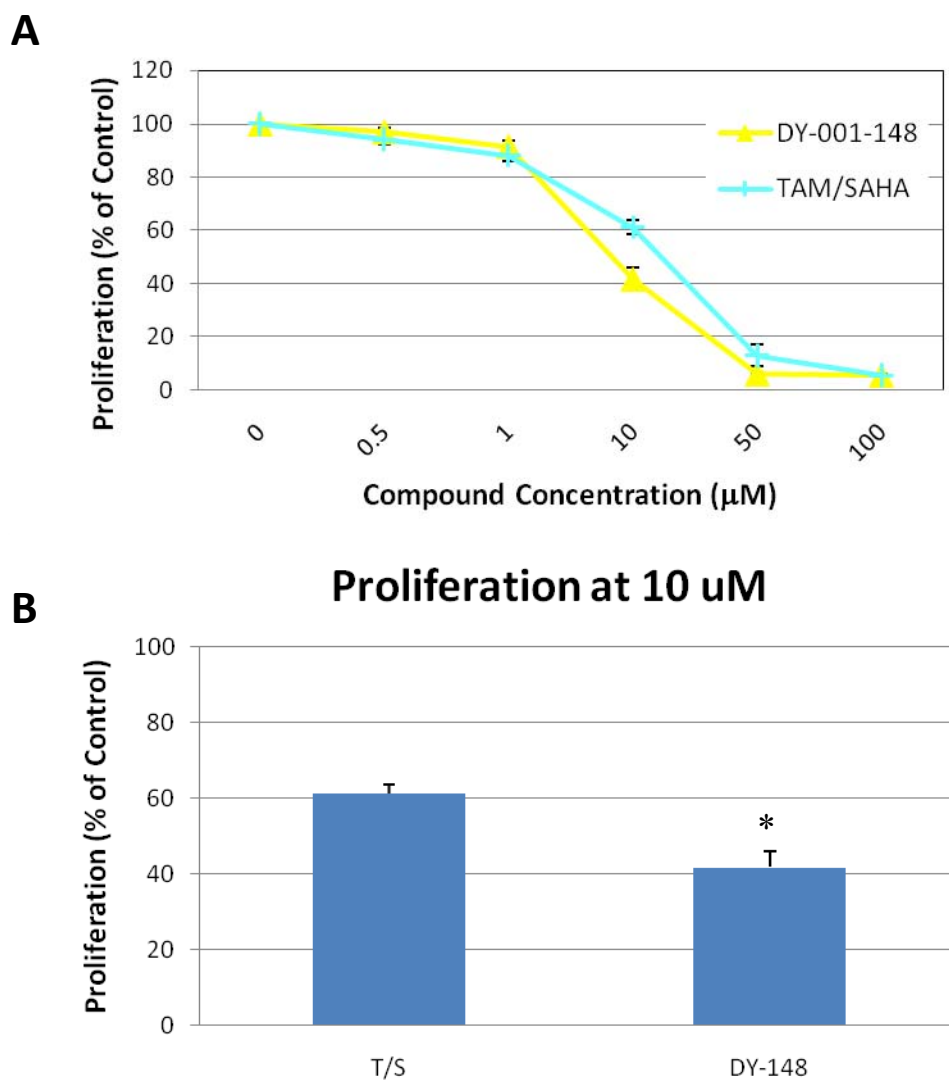


Figure 6.12: Proliferation of MCF-7 Cells in the presence of DY-001-148 and Equal amounts of Tam and SAHA. (A) Dose Response Curves and (B) Proliferation at 10 μ M ligand. Results correspond to means and \pm S.E. of three experiments done in triplicate. * $P < 0.005$ vs. equal amounts of Tam/SAHA determined by t-test

moieties, the SAHA-like moiety in DY-001-148 is actually the same structure as SAHA itself.

6.4.2 Effects of Dual Inhibiting Compounds on the Viability of ER α Negative Breast Cancer Cells

The MDA-MB-231 cell line is a highly invasive breast cancer cell line and is a prototype for the study of hormone-independent breast cancer due to its low ER expression levels [73]. In MDA-MB-231 cells, DY-001-137 and DY-001-138 show approximately the same potency as tamoxifen (Table 6.1). DY-001-148, IC₅₀ value of 8 μ M, is observed to be more potent than tamoxifen, IC₅₀ value of 24 μ M, in MDA-MB-231 cells (Table 6.1). Furthermore, as expected, at 10 μ M concentrations DY-001-148 has an enhanced efficacy in comparison to tamoxifen with a decrease in cell proliferation of 45% versus no decrease with tamoxifen (Figure 6.13).

In comparing the three dual inhibiting compounds with SAHA in MDA-MB-231 cells, both DY-001-137 and DY-001-138 are less potent and have a decreased efficacy in comparison to the SAHA, probably due to the difference in the structures of the SAHA moieties. DY-001-148, however, is more potent than SAHA in MDA-MB-231 cells, with IC₅₀ values of 8 μ M vs. 16 μ M, respectively. Furthermore, at 50 μ M concentrations, DY-001-148 has an enhanced efficacy as compared to SAHA in these cells, with a 90% vs. 70% decrease in cell proliferation (Figure 6.14 & Table 6.1). When evaluating the effects of covalently linking the SAHA-like and tamoxifen-like moieties, in MDA-MB-231 cells, both DY-001-137 and DY-001-138 are less potent (IC₅₀s of 17 μ M and 14 μ M, respectively) and have a decreased efficacy (15% and 20% decrease in cell proliferation, respectively, at 10 μ M) in comparison to adding equal amounts of tamoxifen and SAHA

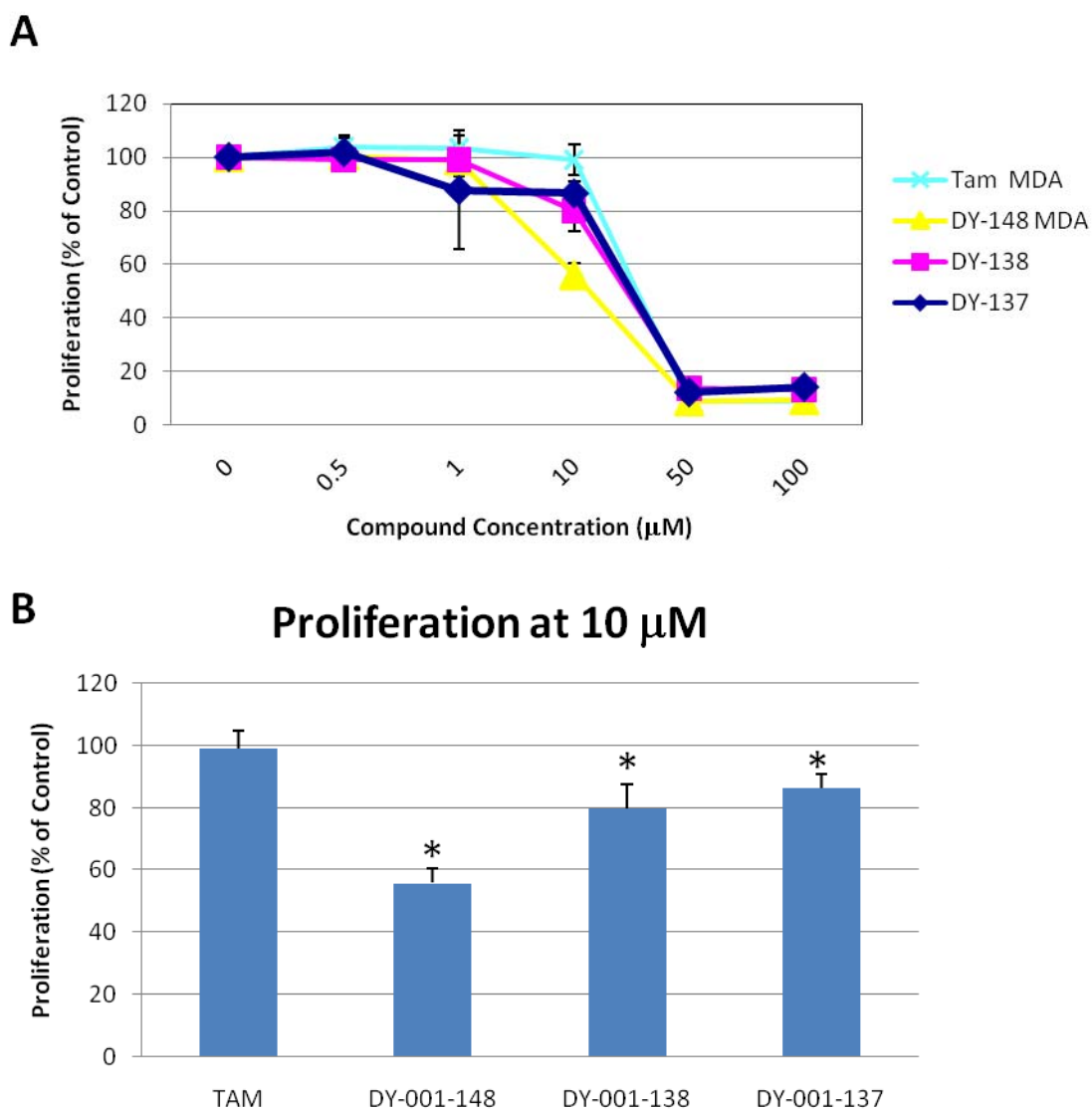


Figure 6.13: Proliferation of MDA-MB-231 Cells in the presence of Tamoxifen, DY-001-137, DY-001-138, and DY-001-148. (A) Dose Response Curves and (B) Proliferation at 10 μM ligand. Results correspond to means and \pm S.D. of two to three experiments done in triplicate. * $P < 0.005$ vs. Tam alone determined by t-test

non-covalently linked, (IC_{50} of 7 μ M and 50% decrease in cell proliferation at 10 μ M).

DY-001-148 has roughly the same potency and efficacy as non-covalently linked tamoxifen and SAHA (Figure 6.15), indicating that the covalent linkage does not have a significant effect on activity in the ER negative breast cancer cell line, MDA-MB-231.

6.4.3 Conclusions of Effects of Dual Inhibiting Compounds in Breast Cancer Cells

Whole cell proliferation assays have confirmed that of all the dual inhibiting compounds tested, DY-001-148 is the most effective of the compounds and is more potent than both tamoxifen alone and SAHA alone in both cell lines. The enhanced efficacy and potency of DY-001-148 in comparison to DY-001-137 and DY-001-138, is most likely a result of the SAHA moiety on DY-001-148 being a more effective HDACi than the SAHA derivatives in the DY-001-137 and DY-001-138 compounds. At 10 μ M, DY-001-148 has an enhanced efficacy in comparison to tamoxifen in both cell lines and an enhanced efficacy in comparison to SAHA in MCF-7 cells. All of these results suggest that the anti-proliferative effects observed are not a result of one moiety's activity over the other, but both moieties in the compound are responsible for the decrease in proliferation observed. DY-001-148 also has an enhanced efficacy in MCF-7 cells in comparison to non-covalently linked tamoxifen and SAHA, suggesting that covalently linking the two compounds increases its effectiveness in MCF-7 ER⁺ breast cancer cells. The fact that the covalent linkage was found to increase efficacies and potencies in ER⁺ breast cancer cells and not ER⁻ breast cancer cells supports the idea that the tamoxifen moiety is sequestering the dual inhibiting compounds into the ER⁺ cells resultantly increasing the effectiveness of the HDACi in these cells. This phenomenon is most likely not seen in ER⁻ breast cancer cells because there is not an overexpression of ER α to

attract the tamoxifen based moiety that would resultantly sequester the dual inhibiting compounds into the cells.

Interestingly, tamoxifen was found to have the same IC_{50} values in both MCF-7 and MDA-MB-231 cells. The IC_{50} value we calculated for tamoxifen in MCF-7 cells is consistent with that found in literature [74-76]. Generally, MDA-MB-231 cells, are supposed to be unresponsive to tamoxifen, however researchers have found that MDA-MB-231 cells are sensitive to supra-physiological concentrations of tamoxifen [77-79]. Groleau *et al.* reported IC_{50} values for tamoxifen, incubated for 72 hrs, in MCF-7 and MDA-MB-231 cells to be 16 μ M and 40 μ M respectively when using a MTT proliferation assay which are comparable to our values [74].

To support the HDACi activity we observed in proliferation assays, the HDAC inhibition activity of the dual inhibiting compounds was tested by the Oyelere lab using the Fluor de Lys assay. The Fluor de Lys assay measures the deacetylation of a substrate using fluorescence. Essentially, if the substrate is deacetylated a substrate fluorophore is generated. The ability of an HDACi to inhibit the deacetylation of the substrate by HDACs can be measure as a decrease in fluorescence. The Oyelere lab found that the dual inhibiting compounds do not yield as low a K_i as SAHA but do show HDAC inhibition (K_i of DY-001-148 was approximately 350 nM and 57 nM with SAHA, data not shown).

6.5 Summary of Dual Inhibiting Compounds as Breast Cancer Therapeutics

Several laboratories have reported that varying HDACi enhance the effects of tamoxifen in ER⁺ breast cancer cell lines and sensitize tamoxifen resistant or ER⁻ breast cancer cell lines to tamoxifen [57-65]. We hypothesize that the dual inhibiting

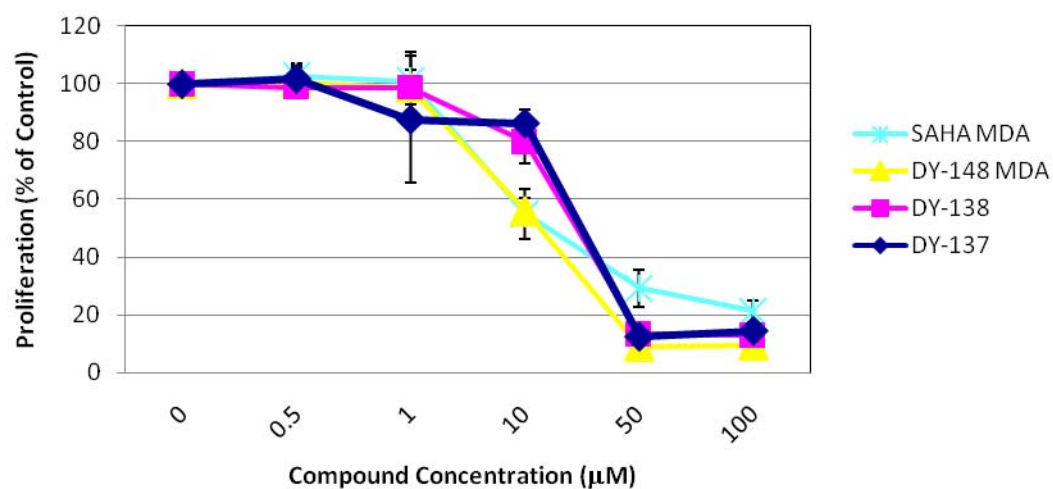
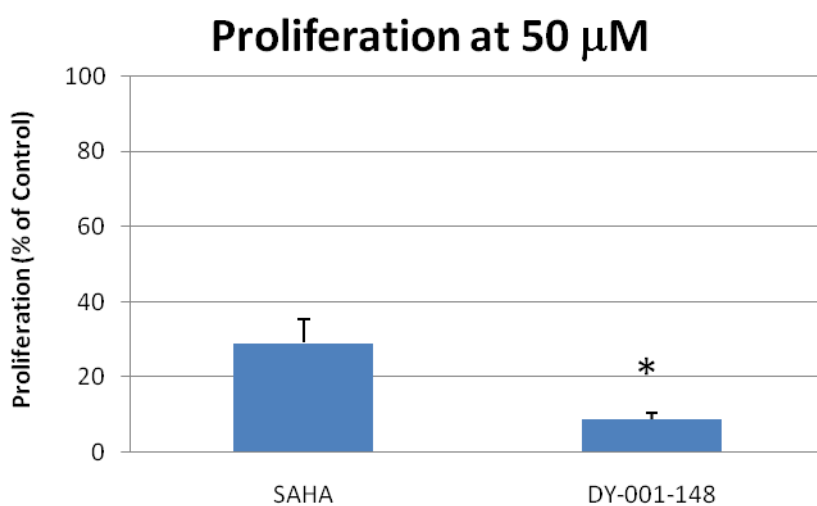
A**B**

Figure 6.14: Proliferation of MDA-MB-231 Cells in the presence of SAHA, DY-001-137, DY-001-138, and DY-001-148. (A) Dose Response Curves and (B) Proliferation at 50 μM ligand. Results correspond to means and \pm S.D. of three experiments done in triplicate. * $P < 0.005$ vs. SAHA only determined by t-test

compounds studied, which contain both HDACi-like and SERM-like moieties, could not only potentially affect both ER⁺ and ER⁻ breast cancers, based on previous work done with non-covalently linked HDACi and tamoxifen [57-65] but, due to the covalent linkage, these compounds could potentially selectively target ER⁺ breast cancers producing an enhanced breast cancer therapeutic than can treat over 70% of all breast cancers. These compounds were found to not only successfully antagonize the effects of estrogen receptor and inducing anti-proliferative effects in both ER⁺ and ER⁻ breast cancer cells. Results with CC show that all three dual inhibiting compounds tested, DY-001-137, DY-001-138, and DY-001-148, are not only capable of antagonizing the effects of estradiol in ER α and β , but are more effective than tamoxifen at doing so. Whole cell proliferation assays corroborate the data obtained in yeast and reveal that we have created a single compound, DY-001-148, with anti-proliferative effects in both ER⁺ and ER⁻ breast cancer cells lines and with a higher potency than with either tamoxifen or SAHA alone in ER⁺ cell lines. We have also established that covalent linkage of SAHA and tamoxifen enhances the anti-proliferative effects of this compounds in MCF-7 cells.

6.6 Future Work

The dual inhibiting compounds tested serve as a starting point for creating other dual inhibiting compounds by covalently linking other HDACi-like moieties, such as Trichostatin A (TSA), to other known SERMs, such as raloxifene or ICI 164384. Lastly, it would also be interesting to investigate how the expression levels of proteins known to be affected by HDACi, such as p21 and acetylated histones (H3 and H4), are affected by the dual inhibiting compounds. We are also testing the effects of these compounds in non

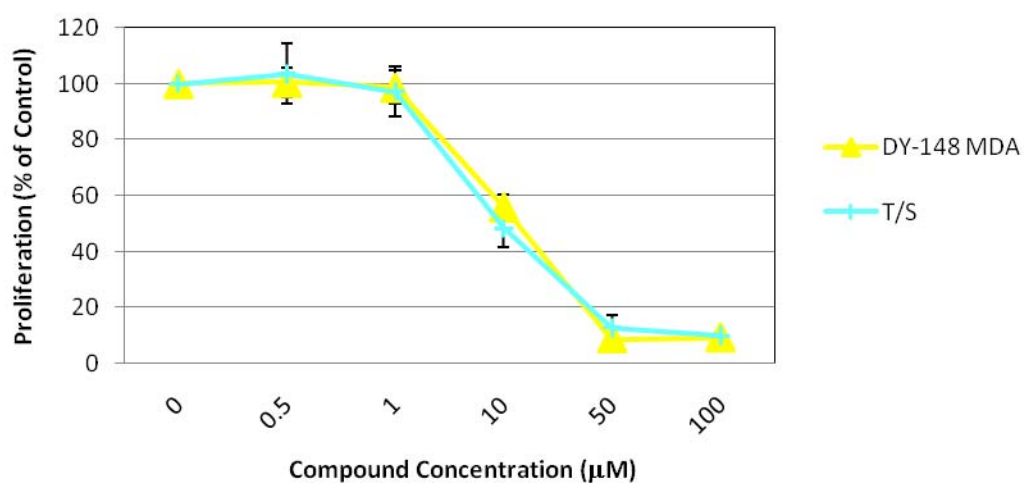


Figure 6.15: Proliferation of MDA-MB-231 Cells in the presence of DY-001-148 and Equal amounts of Tam and SAHA. Results correspond to means and \pm S.D. of three experiments done in triplicate. * $P < 0.005$ vs. equal amounts of Tam/SAHA determined by t-test

ER overexpression cells to determine whether the covalent linkage of the HDACi to tamoxifen is truly capable of selectively targeting ER⁺ cells.

6.7 Materials and Methods

Yeast Strain

Yeast strain, PJ69-4A (*MATa trp1-901 leu2-3, 112 ura3-52 his3-200 gal4Δ gal80Δ LYS2::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ*), was a kind gift from Dr. Philip James and Dr. Elizabeth Craig (University of Wisconsin, Madison) [80]. The strain contains the *HIS3*, *ADE2*, and *lacZ* genes each under the control of Gal4 promoter.

Ligands

17β-estradiol (MW=272.4 g/mol) was purchased from MP Biomedicals, (Salon, OH). Tamoxifen (MW=371.5 g/mol) was purchased from Sigma-Aldrich (St. Louis, MO). Suberoylanilide hydroxamic acid (SAHA), DY-001-135, DY-001-136, DY-001-137, DY-001-138, DY-001-148, and AO-002-004 were synthesized by Dr. Oyelere's Lab (Georgia Institute of Technology, Georgia). 10 mM stocks of the ligand were dissolved in 80% ethanol:20% DMSO (4:1 v/v) and stored at 4°C.

Expression Plasmids

The pGAD10BA ACTR vector, containing a tryptophan marker, were previously constructed in our lab and contain a Gal4AD fused to the full coactivator, ACTR [71, 81]. To make pGBDERαLBD, containing a leucine marker, primers containing a NheI site at the 5' end and SpeI site at the 3' end were designed to amplify the ERαLBD fusion gene (containing residues 301-595 of ER α) from the pSG5-HEGO vector. Once amplified,

the gene was ligated into the pGBDRXR vector using the NheI and SpeI restriction enzymes.

Yeast Transformations and Selective Media and Plates

Synthetic complete (SC) plates were made as described previously [82]. Selective media and plates were made of SC media minus, either, leucine and tryptophan or minus histidine, leucine, and tryptophan. pGBDER α LBD and either pGAD10BA ACTR , pGAD10BA SRC1, or pGAD10BA- mPGC1a were introduced into PJ69-4A using the LiAc transformation method [83]. Cells were plated onto SC –Leu-Trp plates and inoculated.

Yeast ER α Agonist and Antagonist Assays

Yeast quantitation and β -galactosidase assays were performed as previously described [81]. Briefly, in the growth assay, cells were grown in SC-His-Leu-Trp media with the appropriate ligands in of 96-well microtiter plate at 30 °C and 150 RPM. OD₆₃₀ was measured at t=0, t=24 hours, and t=48 hours to determine growth in Media lacking histidine. B-galactosidase assays were performed by growing cells in SC-Leu-Trp media with the appropriate ligands in of 96-well microtiter plate at 30 °C and 150 RPM. OD₆₃₀ was measured at t=0, t=24 hours, and t=48 hours. After 48 hours, cells were lysed, ONPG was added, and once media in wells became yellow, OD₄₀₅ was measured to quantify β -galactosidase activity.

Cell Culture and MTS Cell Proliferation Assay

MCF-7 and MDA-MB-231 cells were provided by Dr. Al Merrill (Georgia Institute of Technology, GA) and were routinely cultured in DMEM (Invitrogen, Grand Island, NY) with 10% fetal bovine serum (FBS) (Hycone, Logan, UT) and antibiotics.

For all experiments cells were grown in 96-well cell culture treated microtiter plates (Corning Inc., Corning, NY) with the appropriate ligand in triplicate for 72 hours. MTS, CellTiter 96 Aqueous One Solution Cell Proliferation Assays (Promega, Madison, WI), were performed according to manufacturer's instructions

6.8 Literature Cited

1. Shiau AK, Barstad D, Loria PM, Cheng L, Kushner PJ, Agard DA & Greene GL (1998) The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* **95**, 927-937.
2. Soerjomataram I, Pukkala E, Brenner H & Coebergh JW (2008) On the avoidability of breast cancer in industrialized societies: older mean age at first birth as an indicator of excess breast cancer risk. *Breast Cancer Res Treat* **111**, 297-302.
3. Kelsey JL & Berkowitz GS (1988) Breast cancer epidemiology. *Cancer Res* **48**, 5615-5623.
4. Deroo BJ & Korach KS (2006) Estrogen receptors and human disease. *J Clin Invest* **116**, 561-570.
5. Arpino G, Weiss H, Lee AV, Schiff R, De Placido S, Osborne CK & Elledge RM (2005) Estrogen receptor-positive, progesterone receptor-negative breast cancer: association with growth factor receptor expression and tamoxifen resistance. *J Natl Cancer Inst* **97**, 1254-1261.
6. Yue W, Wang JP, Li Y, Bocchinfuso WP, Korach KS, Devanesan PD, Rogan E, Cavalieri E & Santen RJ (2005) Tamoxifen versus aromatase inhibitors for breast cancer prevention. *Clin Cancer Res* **11**, 925s-930s.
7. Imamov O, Morani A, Shim GJ, Omoto Y, Thulin-Andersson C, Warner M & Gustafsson JA (2004) Estrogen receptor beta regulates epithelial cellular differentiation in the mouse ventral prostate. *Proc Natl Acad Sci U S A* **101**, 9375-9380.
8. Hedden A, Muller V & Jensen EV (1995) A new interpretation of antiestrogen action. *Ann N Y Acad Sci* **761**, 109-120.
9. Jordan VC (2001) The past, present, and future of selective estrogen receptor modulation. *Ann N Y Acad Sci* **949**, 72-79.
10. Park WC & Jordan VC (2002) Selective estrogen receptor modulators (SERMS) and their roles in breast cancer prevention. *Trends Mol Med* **8**, 82-88.

11. Bjornstrom L & Sjoberg M (2005) Mechanisms of estrogen receptor signaling: convergence of genomic and nongenomic actions on target genes. *Mol Endocrinol* **19**, 833-842.
12. Kang KI, Devin J, Cadepond F, Jibard N, Guiochon-Mantel A, Baulieu EE & Catelli MG (1994) In vivo functional protein-protein interaction: nuclear targeted hsp90 shifts cytoplasmic steroid receptor mutants into the nucleus. *Proc Natl Acad Sci U S A* **91**, 340-344.
13. Li X, Huang J, Yi P, Bambara RA, Hilf R & Muyan M (2004) Single-chain estrogen receptors (ERs) reveal that the ERalpha/beta heterodimer emulates functions of the ERalpha dimer in genomic estrogen signaling pathways. *Mol Cell Biol* **24**, 7681-7694.
14. Ascenzi P, Bocedi A & Marino M (2006) Structure-function relationship of estrogen receptor alpha and beta: impact on human health. *Mol Aspects Med* **27**, 299-402.
15. Bourguet W, Germain P & Gronemeyer H (2000) Nuclear receptor ligand-binding domains: three-dimensional structures, molecular interactions and pharmacological implications. *Trends Pharmacol Sci* **21**, 381-388.
16. Cheskis BJ (2004) Regulation of cell signalling cascades by steroid hormones. *J Cell Biochem* **93**, 20-27.
17. Levin ER (2005) Integration of the extranuclear and nuclear actions of estrogen. *Mol Endocrinol* **19**, 1951-1959.
18. Shupnik MA (2004) Crosstalk between steroid receptors and the c-Src-receptor tyrosine kinase pathways: implications for cell proliferation. *Oncogene* **23**, 7979-7989.
19. Razandi M, Pedram A & Levin ER (2000) Plasma membrane estrogen receptors signal to antiapoptosis in breast cancer. *Mol Endocrinol* **14**, 1434-1447.
20. Kousteni S, Han L, Chen JR, Almeida M, Plotkin LI, Bellido T & Manolagas SC (2003) Kinase-mediated regulation of common transcription factors accounts for the bone-protective effects of sex steroids. *J Clin Invest* **111**, 1651-1664.
21. Razandi M, Alton G, Pedram A, Ghonshani S, Webb P & Levin ER (2003) Identification of a structural determinant necessary for the localization and function of estrogen receptor alpha at the plasma membrane. *Mol Cell Biol* **23**, 1633-1646.
22. Razandi M, Pedram A, Park ST & Levin ER (2003) Proximal events in signaling by plasma membrane estrogen receptors. *J Biol Chem* **278**, 2701-2712.
23. Song RX, Fan P, Yue W, Chen Y & Santen RJ (2006) Role of receptor complexes in the extranuclear actions of estrogen receptor alpha in breast cancer. *Endocr Relat Cancer* **13 Suppl 1**, S3-13.

24. O'Malley BW (2005) A life-long search for the molecular pathways of steroid hormone action. *Mol Endocrinol* **19**, 1402-1411.
25. Kansra S, Yamagata S, Sneade L, Foster L & Ben-Jonathan N (2005) Differential effects of estrogen receptor antagonists on pituitary lactotroph proliferation and prolactin release. *Mol Cell Endocrinol* **239**, 27-36.
26. Shang Y & Brown M (2002) Molecular determinants for the tissue specificity of SERMs. *Science* **295**, 2465-2468.
27. Love RR, Mazess RB, Barden HS, Epstein S, Newcomb PA, Jordan VC, Carbone PP & DeMets DL (1992) Effects of tamoxifen on bone mineral density in postmenopausal women with breast cancer. *N Engl J Med* **326**, 852-856.
28. Ismail SM (1994) Effects of tamoxifen on uterus. *Lancet* **344**, 622-623.
29. Howell A, Cuzick J, Baum M, Buzdar A, Dowsett M, Forbes JF, Hocht-Boes G, Houghton J, Locker GY & Tobias JS (2005) Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer. *Lancet* **365**, 60-62.
30. Jordan VC (2006) Tamoxifen (ICI46,474) as a targeted therapy to treat and prevent breast cancer. *Br J Pharmacol* **147 Suppl 1**, S269-276.
31. Legha SS, Davis HL & Muggia FM (1978) Hormonal therapy of breast cancer: new approaches and concepts. *Ann Intern Med* **88**, 69-77.
32. Sutherland RL, Green MD, Hall RE, Reddel RR & Taylor IW (1983) Tamoxifen induces accumulation of MCF 7 human mammary carcinoma cells in the G0/G1 phase of the cell cycle. *Eur J Cancer Clin Oncol* **19**, 615-621.
33. Mikulski SM, Viera A, Ardelt W, Menduke H & Shogen K (1990) Tamoxifen and trifluoroperazine (Stelazine) potentiate cytostatic/cytotoxic effects of P-30 protein, a novel protein possessing anti-tumor activity. *Cell Tissue Kinet* **23**, 237-246.
34. Grilli S (2006) Tamoxifen (TAM): the dispute goes on. *Ann Ist Super Sanita* **42**, 170-173.
35. Nakamura T, Imai Y, Matsumoto T, Sato S, Takeuchi K, Igarashi K, Harada Y, Azuma Y, Krust A, Yamamoto Y, Nishina H, Takeda S, Takayanagi H, Metzger D, Kanno J, Takaoka K, Martin TJ, Chambon P & Kato S (2007) Estrogen prevents bone loss via estrogen receptor alpha and induction of Fas ligand in osteoclasts. *Cell* **130**, 811-823.
36. Krum SA, Miranda-Carboni GA, Hauschka PV, Carroll JS, Lane TF, Freedman LP & Brown M (2008) Estrogen protects bone by inducing Fas ligand in osteoblasts to regulate osteoclast survival. *EMBO J* **27**, 535-545.

37. Rochefort H, Glondou M, Sahla ME, Platet N & Garcia M (2003) How to target estrogen receptor-negative breast cancer? *Endocr Relat Cancer* **10**, 261-266.
38. Zhang Z, Yamashita H, Toyama T, Sugiura H, Ando Y, Mita K, Hamaguchi M, Hara Y, Kobayashi S & Iwase H (2005) Quantitation of HDAC1 mRNA expression in invasive carcinoma of the breast*. *Breast Cancer Res Treat* **94**, 11-16.
39. Brehm A, Miska EA, McCance DJ, Reid JL, Bannister AJ & Kouzarides T (1998) Retinoblastoma protein recruits histone deacetylase to repress transcription. *Nature* **391**, 597-601.
40. Attar E & Bulun SE (2006) Aromatase inhibitors: the next generation of therapeutics for endometriosis? *Fertil Steril* **85**, 1307-1318.
41. Mokbel K (2002) The evolving role of aromatase inhibitors in breast cancer. *Int J Clin Oncol* **7**, 279-283.
42. Sanderson JT, Hordijk J, Denison MS, Springsteel MF, Nantz MH & van den Berg M (2004) Induction and inhibition of aromatase (CYP19) activity by natural and synthetic flavonoid compounds in H295R human adrenocortical carcinoma cells. *Toxicol Sci* **82**, 70-79.
43. Martinez-Iglesias O, Ruiz-Llorente L, Sanchez-Martinez R, Garcia L, Zambrano A & Aranda A (2008) Histone deacetylase inhibitors: mechanism of action and therapeutic use in cancer. *Clin Transl Oncol* **10**, 395-398.
44. Munster PN, Troso-Sandoval T, Rosen N, Rifkind R, Marks PA & Richon VM (2001) The histone deacetylase inhibitor suberoylanilide hydroxamic acid induces differentiation of human breast cancer cells. *Cancer Res* **61**, 8492-8497.
45. Marks PA, Richon VM & Rifkind RA (2000) Histone deacetylase inhibitors: inducers of differentiation or apoptosis of transformed cells. *J Natl Cancer Inst* **92**, 1210-1216.
46. Guenther MG, Barak O & Lazar MA (2001) The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3. *Mol Cell Biol* **21**, 6091-6101.
47. Love JD, Gooch JT, Nagy L, Chatterjee VK & Schwabe JW (2000) Transcriptional repression by nuclear receptors: mechanisms and role in disease. *Biochem Soc Trans* **28**, 390-396.
48. Nagy L, Kao HY, Chakravarti D, Lin RJ, Hassig CA, Ayer DE, Schreiber SL & Evans RM (1997) Nuclear receptor repression mediated by a complex containing SMRT, mSin3A, and histone deacetylase. *Cell* **89**, 373-380.
49. Nagy L, Kao HY, Love JD, Li C, Banayo E, Gooch JT, Krishna V, Chatterjee K, Evans RM & Schwabe JW (1999) Mechanism of corepressor binding and release from nuclear hormone receptors. *Genes Dev* **13**, 3209-3216.

50. Glozak MA, Sengupta N, Zhang X & Seto E (2005) Acetylation and deacetylation of non-histone proteins. *Gene* **363**, 15-23.
51. (1985) Chemotherapy with or without oophorectomy in high-risk premenopausal patients with operable breast cancer. *J Clin Oncol* **3**, 1059-1067.
52. Clarke MJ (2008) WITHDRAWN: Tamoxifen for early breast cancer. *Cochrane Database Syst Rev*, CD000486.
53. Bardou VJ, Arpino G, Elledge RM, Osborne CK & Clark GM (2003) Progesterone receptor status significantly improves outcome prediction over estrogen receptor status alone for adjuvant endocrine therapy in two large breast cancer databases. *J Clin Oncol* **21**, 1973-1979.
54. Olsen Eea (2006) Vorinostat (suberoylanilide hydroxamic acid, SAHA) is clinically active in advanced cutaneous T-cell Lymphoma (CTCL): results of Phase IIB trial. In *ASCO Annual Meeting Proceedings Part I*, pp. 7500.
55. Frew AJ, Lindemann RK, Martin BP, Clarke CJ, Sharkey J, Anthony DA, Banks KM, Haynes NM, Gangatirkar P, Stanley K, Bolden JE, Takeda K, Yagita H, Secrist JP, Smyth MJ & Johnstone RW (2008) Combination therapy of established cancer using a histone deacetylase inhibitor and a TRAIL receptor agonist. *Proc Natl Acad Sci U S A* **105**, 11317-11322.
56. Thomas S & Munster PN (2009) Histone deacetylase inhibitor induced modulation of anti-estrogen therapy. *Cancer Lett*, [Epub ahead of print]
57. Hodges-Gallagher LV, C; El Bader, S; Kushner, P (2007) Inhibition of histone deacetylase enhance the anti-proliferative action of antiestrogens on breast cancer cells and blocks tamoxifen-induced proliferation of uterine cells. *Breast Cancer Research and Treatment* **105**, 297-309.
58. Jang ER LS, Lee ES, Jeong G, Kim TY, Bang YJ, Lee JS (2004) The histone deacetylase inhibitor trichostatin A sensitizes estrogen receptor alpha-negative breast cancer cells to tamoxifen. *Oncogene* **23**, 1724-1736.
59. Margueron RD, V; Bonnet, S; Escande, A; Vignon, F; Balaguer, P; Cavailles, V (2004) Histone deacetylase inhibition and estrogen receptor alpha levels modulate the transcription activity of partial antiestrogens. *Journal of Molecular Endocrinology* **32**, 583-594.
60. Sharma DS, N; Davidson, N; Vertino, P (2006) Restoration of Tamoxifen Sensitivity in Estrogen Receptor-Negative Breast Cancer Cells: Tamoxifen-Bound Reactivated ER Recruits Distinctive Corepressor Complexes. *Cancer Research* **66**, 6370-6378.
61. Saji S, Kawakami M, Hayashi S, Yoshida N, Hirose M, Horiguchi S, Itoh A, Funata N, Schreiber SL, Yoshida M & Toi M (2005) Significance of HDAC6 regulation

via estrogen signaling for cell motility and prognosis in estrogen receptor-positive breast cancer. *Oncogene* **24**, 4531-4539.

62. Fan J, Yin WJ, Lu JS, Wang L, Wu J, Wu FY, Di GH, Shen ZZ & Shao ZM (2008) ER alpha negative breast cancer cells restore response to endocrine therapy by combination treatment with both HDAC inhibitor and DNMT inhibitor. *J Cancer Res Clin Oncol* **134**, 883-890.

63. Fiskus W, Ren Y, Mohapatra A, Bali P, Mandawat A, Rao R, Herger B, Yang Y, Atadja P, Wu J & Bhalla K (2007) Hydroxamic acid analogue histone deacetylase inhibitors attenuate estrogen receptor-alpha levels and transcriptional activity: a result of hyperacetylation and inhibition of chaperone function of heat shock protein 90. *Clin Cancer Res* **13**, 4882-4890.

64. Hirokawa Y, Arnold M, Nakajima H, Zalcberg J & Maruta H (2005) Signal therapy of breast cancers by the HDAC inhibitor FK228 that blocks the activation of PAK1 and abrogates the tamoxifen-resistance. *Cancer Biol Ther* **4**, 956-960.

65. Liu XF & Bagchi MK (2004) Recruitment of distinct chromatin-modifying complexes by tamoxifen-complexed estrogen receptor at natural target gene promoters in vivo. *J Biol Chem* **279**, 15050-15058.

66. Piekarz RL, Frye AR, Wright JJ, Steinberg SM, Liewehr DJ, Rosing DR, Sachdev V, Fojo T & Bates SE (2006) Cardiac studies in patients treated with depsipeptide, FK228, in a phase II trial for T-cell lymphoma. *Clin Cancer Res* **12**, 3762-3773.

67. Shah MH, Binkley P, Chan K, Xiao J, Arbogast D, Collamore M, Farra Y, Young D & Grever M (2006) Cardiotoxicity of histone deacetylase inhibitor depsipeptide in patients with metastatic neuroendocrine tumors. *Clin Cancer Res* **12**, 3997-4003.

68. Molife R, Fong P, Scurr M, Judson I, Kaye S & de Bono J (2007) HDAC inhibitors and cardiac safety. *Clin Cancer Res* **13**, 1068; author reply 1068-1069.

69. Kell J (2007) Drug evaluation: MGCD-0103, a histone deacetylase inhibitor for the treatment of cancer. *Curr Opin Investig Drugs* **8**, 485-492.

70. Azizi B, Chang EI & Doyle DF (2003) Chemical complementation: small-molecule-based genetic selection in yeast. *Biochem Biophys Res Commun* **306**, 774-780.

71. Azizi B, Chang EI, Schwimmer LJ & Doyle DF (2009) *In Progress*.

72. Schwimmer LJ, Rohatgi P, Azizi B, Seley KL & Doyle DF (2004) Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules. *Proc Natl Acad Sci U S A* **101**, 14707-14712.

73. Ramaswamy S (2006) Cancer transcriptomics: modeling metastasis. *Eur J Hum Genet* **14**, 1-2.

74. Groleau S, Nault, J, Lepage, M, Couture, M, Dallaire, N, Be'rube', G, and C.-Gaudreault, R (1999) Synthesis and Preliminary *in Vitro* Cytotoxic Activity of New Triphenylethylene Dimers. *Bioorganic Chemistry* **27**, 383-394.
75. Tari AM, Simeone AM, Li YJ, Gutierrez-Puente Y, Lai S & Symmans WF (2005) Cyclooxygenase-2 protein reduces tamoxifen and N-(4-hydroxyphenyl)retinamide inhibitory effects in breast cancer cells. *Lab Invest* **85**, 1357-1367.
76. Wilson ST, Blask DE & Lemus-Wilson AM (1992) Melatonin augments the sensitivity of MCF-7 human breast cancer cells to tamoxifen in vitro. *J Clin Endocrinol Metab* **75**, 669-670.
77. Perry RR, Kang Y & Greaves BR (1995) Relationship between tamoxifen-induced transforming growth factor beta 1 expression, cytostasis and apoptosis in human breast cancer cells. *Br J Cancer* **72**, 1441-1446.
78. Rosfjord EC & Dickson RB (1999) Growth factors, apoptosis, and survival of mammary epithelial cells. *J Mammary Gland Biol Neoplasia* **4**, 229-237.
79. Berube G, Rabouin D, Perron V, N'Zemba B, Gaudreault RC, Parent S & Asselin E (2006) Synthesis of unique 17beta-estradiol homo-dimers, estrogen receptors binding affinity evaluation and cytotoxic activity on breast, intestinal and skin cancer cell lines. *Steroids* **71**, 911-921.
80. James P, Halladay J & Craig EA (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144**, 1425-1436.
81. Azizi B (2005) Chemical complementation a genetic selection system in yeast for drug discovery, protein engineering, and for deciphering and assembling biosynthetic pathways. In *unpublished thesis (PhD)*. Georgia Institute of Technology, Atlanta.
82. Frederick M. Ausubel RB, Robert E. Kingston, David D. Moore, J. G. Seidman, John A. Smith, Kevin Struhl (1999) *Short Protocols in Molecular Biology*, 4th edn. Wiley John & Sons Inc.
83. Guthrie C (2002) Guide to Yeast Genetics and Molecular Cell Biology, Part B In *Methods in Enzymology*. Academic Press.